

SEARCH REQUEST FORM

03-1055

Examiner # (Mandatory): _____ Requester's Full Name: Andrew Wong
 Art Unit 1635 Location (Bldg/Room#): CM 111009 Phone (circle 305/306/308) 3217
 Serial Number: 091238,972 Results Format Preferred (circle): PAPER DISK E-MAIL

Title of Invention _____

Inventors (please provide full names): _____

Earliest Priority Date: _____

Keywords (include any known synonyms registry numbers, explanation of initialisms):

*Andrew,
Seq. ID 2 Searched w/*

*Andrew,
Seq. ID 2 Searched w/ & w/o bp
limits.
Beverly*

Search Topic:

Please write detailed statement of the search topic, and the concept of the invention. Describe as specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples of relevant citations, authors, etc., if known. You may include a copy of the abstract and the broadcast or most relevant claim(s).

Please search SEQ ID No. 2

PLEASE RUSH.

Moye Elliott

*thanks
Drew*

Point of Contact:
Beverly Shears
Technical Info. Specialist
CM1 12C14 Tel: 308-4994

STAFF USE ONLY

Searcher: Beverly 24994 Type of Search _____ Vendors (include cost where applicable)
 Searcher Phone #: _____ N.A. Sequence ☒ STN
 Searcher Location: _____ A.A. Sequence _____ Questel/Orbit
 Date Picked Up: _____ Structure (#) _____ Lexis/Nexis
 Date Completed: 04-02-99 Bibliographic _____ WWW/Internet
 Clerical Prep Time: 3 Litigation I _____ In-house sequence systems (list)
 Terminal Time: 116 Fulltext _____ Dialog
 Number of Databases: 2 Procurement _____ Dr. Link
 _____ Other _____ Westlaw
 _____ Other (specify) mb

THIS PAGE BLANK (USPTO)

09/238972

09/238972

FILE 'REGISTRY' ENTERED AT 11:29:26 ON 02 APR 1999
L18 0 S GTAGGCTGAAACCCTGTCCTTGC/SQSN

=> fil hom

FILE 'HOME' ENTERED AT 11:30:12 ON 02 APR 1999

Searcher : Shears 308-4994

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)


```

DE Oligonucleotide probe MK14-A
KW Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
KW ss.
OS Synthetic.
PN EP-571911-A.
PD 01-DEC-1993.
PF 24-MAY-1993; 108325.
PR 26-MAY-1992; US-889651.
PA (BECT ) BECTON DICKINSON CO.
PI Shank DD, Spears PA;
DR WPI; 93-37884/48.
PT New oligo:nucleotide probes specific for Mycobacteria - used for
PT detection and amplification of Mycobacteria nucleic acid in
PT samples
PS Claim 3; Page 14; 23pp; English.
CC Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
CC (Q51735). It hybridized to all spp. of Mycobacteria tested, but
CC cross reacted to a few non-mycobacterial spp. The probe may
CC be useful as an initial screen for mycobacterial infection.
CC See also Q51735-45 and Q51747-59.
SQ Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;

Query Match 78.3%; Score 18; DB 9; Length 91;
Best Local Similarity 0.0%; Pred. No. 3.68e+01;
Matches 0; Conservative 20; Mismatches 2; Indels 0; Gaps 0;

Db 25 hshvhhvvhhvshvvvhhv 46
QY 2 tagtgtaaacctgtctctgc 23

RESULT 3
ID Q87382 standard; DNA; 25 BP.
AC Q87382;
DT 19-SEP-1995 (first entry)
DE PCR primer 3g (MOG nt1083-1107).
KW MOG; myelin oligodendrocyte glycoprotein; autoimmune disease;
KW multiple sclerosis; anti-idiotype; polymerase chain reaction;
KW PCR; primer; amplification; probe; RNase-H mapping; ss.
OS Synthetic.
PN WO9507096-A.
PD 16-MAR-1995.
PF 02-SEP-1994; AU0522.
PR 06-SEP-1993; AU-001030.
PA (UyLT-) UNIV LA TROBE.
PI Bernard CCA, Kerlero de Rosbo NCM;
DR WPI; 95-123238/16.
PT Treating a T-cell and/or B-cell mediated auto-immune disease - by
PT administering an active agent selected from myelin
PT oligo:dendrocyte protein (MOG), immuno:dominant epitope(s) of MOG
PT or anti-idiotypic antibodies directed against these
PT Disclosure; Page 83; 123pp; English.
PS RNA was purified from the brains of healthy and multiple sclerosis
CC affected individuals. RNase-H digested poly-A RNA was probed with
CC fragments 5' and 3' of the digestion site to determine the size and
CC number of alternative transcripts. The 5' probe was amplified from
CC a lambda gt10 myelin oligodendrocyte glycoprotein (MOG) clone using
CC the primers Ig (given in Q87378) and 6f (Q87377). The primers 9f
CC (Q87379) and 9(3') (Q87380) were used to amplify a 3' probe that
CC excluded a truncated form of MOG, while a probe specific for
CC truncated MOG was amplified using primers 3a (Q87381) and 3g
CC (Q87382).
SQ Sequence 25 BP; 10 A; 5 C; 5 G; 5 T;

Query Match 56.5%; Score 13; DB 14; Length 25;
Best Local Similarity 93.3%; Pred. No. 1.84e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 11 gaaacatgtctctg 25
QY 8 gaaacatgtctctg 22

```

```

RESULT 4
ID T62501 standard; DNA; 27 BP.
AC T62501;
DT 28-APR-1997 (first entry)
DE Murine retrovirus consensus left side integration response sequence.
KW Consensus sequence; left side; integration response; LTR;
KW long terminal repeat; retroviral attachment sequence; preparation;
KW gene delivery construct; ss.
OS Murine retrovirus.
PN WO9626745-A1.
PD 06-SEP-1996.
PF 28-FEB-1996; U02877.
PR 28-FEB-1995; US-395355.
PA (UYCR-) UNIV CREIGHTON.
PI Hodgson CP;
DR WPI; 96-412589/41.
PT Gene delivery system including liposome or dendrimer and
PT perpetuation molecule - esp. for gene therapy, provides efficient
PT and stable expression of chimeric genes more safely than use of
PT viruses.
PS Example 3; Fig 2; 50pp; English.
CC The present sequence is the murine retrovirus consensus left side
CC (U3) integration response sequence, which reads into the long
CC terminal repeat (LTR). It can be used as a retroviral attachment
CC sequence in the preparation of a gene delivery construct.
SQ Sequence 27 BP; 4 A; 0 C; 1 G; 1 T;

Query Match 56.5%; Score 13; DB 28; Length 27;
Best Local Similarity 4.8%; Pred. No. 1.84e+02;
Matches 1; Conservative 14; Mismatches 6; Indels 0; Gaps 0;

Db 6 arrwbvvhvhhndhndknhv 26
QY 3 aggtgaaacctgtctctgc 23

RESULT 5
ID T62501 standard; DNA; 27 BP.
AC T62501;
DT 28-APR-1997 (first entry)
DE Murine retrovirus consensus left side integration response sequence.
KW Consensus sequence; left side; integration response; LTR;
KW long terminal repeat; retroviral attachment sequence; preparation;
KW gene delivery construct; ss.
OS Murine retrovirus.
PN WO9626745-A1.
PD 06-SEP-1996.
PF 28-FEB-1996; U02877.
PR 28-FEB-1995; US-395355.
PA (UYCR-) UNIV CREIGHTON.
PI Hodgson CP;
DR WPI; 96-412589/41.
PT Gene delivery system including liposome or dendrimer and
PT perpetuation molecule - esp. for gene therapy, provides efficient
PT and stable expression of chimeric genes more safely than use of
PT viruses.
PS Example 3; Fig 2; 50pp; English.
CC The present sequence is the murine retrovirus consensus left side
CC (U3) integration response sequence, which reads into the long
CC terminal repeat (LTR). It can be used as a retroviral attachment
CC sequence in the preparation of a gene delivery construct.
SQ Sequence 27 BP; 4 A; 0 C; 1 G; 1 T;

Query Match 56.5%; Score 13; DB 28; Length 27;
Best Local Similarity 0.0%; Pred. No. 1.84e+02;
Matches 0; Conservative 14; Mismatches 5; Indels 0; Gaps 0;

Db 8 rwbvvhvhhndhndknhv 26
QY 21 aaggacagggttcagcct 3

RESULT 6

```


DR WPI: 97-034387/03.
PT Identification of nucleic acid ligands to TGF-beta, PDGF and hKGF
PT using SELEX, used in the diagnosis and treatment of proliferative
PT disorders
PS Example 16; Page 169; 209pp; English.
CC The present sequence is a consensus of human keratinocyte growth
CC factor (hKGF) binding ligands, identified by Systematic Evolution
CC of Ligands by Exponential enrichment (SELEX). Briefly a candidate
CC mixture of nucleic acids was contacted with hKGF, and nucleic
CC acids having an increased affinity to hKGF partitioned from the
CC remainder of the mixture. The partitioned nucleic acids were then
CC amplified to yield a mixture of nucleic acids enriched for
CC sequences with higher affinity and specificity for binding to hKGF.
CC The ligand is anti-mitogenic and may be used to inhibit epithelial
CC cell proliferation, or in the diagnosis and treatment of hKGF
CC mediated pathological conditions, e.g. psoriasis, cancer and
CC inflammatory bowel syndrome.
SQ Sequence 26 BP; 1 A; 0 C; 3 G; 0 U;
Query Match 52.2%; Score 12; DB 31; Length 26;
Best Local Similarity 0.0%; Pred. No. 5.70e+02;
Matches 0; Conservative 15; Mismatches 3; Indels 0; Gaps 0;
Db 9 yysrbsryyybybsy 26
::: :::: ::::: :::::
QY 6 ctgaaccctgtccttgc 23
RESULT 9
ID T62502 standard; DNA; 27 BP.
AC T62502;
DE Murine retrovirus consensus right side integration response sequence.
DE Consensus sequence; right side; integration response; LTR;
KW long terminal repeat; retroviral attachment sequence; preparation;
KW gene delivery construct; ss.
OS Murine retrovirus.
PN WO9626745-A1.
PF 06-SEP-1996.
PF 28-FEB-1996; U02877.
PR 28-FEB-1995; US-395355.
PA (UYCR-) UNIV CREIGHTON.
PI Hodgson CP;
PI WPI: 96-412589/41.
PT Gene delivery system including liposome or dendrimer and
PT perpetuation molecule - esp. for gene therapy, provides efficient
PT and stable expression of chimeric genes more safely than use of
PT viruses.
PS Example 3; Fig 2; 50pp; English.
CC The present sequence is the murine retrovirus consensus right side
CC (U5) integration response sequence, which reads towards the outside
CC of the long terminal repeat (LTR). It can be used as a retroviral
CC attachment sequence in the preparation of a gene delivery construct.
SQ Sequence 27 BP; 1 A; 2 C; 0 G; 6 T;
Query Match 52.2%; Score 12; DB 28; Length 27;
Best Local Similarity 0.0%; Pred. No. 5.70e+02;
Matches 0; Conservative 12; Mismatches 1; Indels 0; Gaps 0;
Db 1 drvbnbsbhbby 13
::: :::: ::::: :::::
QY 9 aaaccctgtcctt 21
RESULT 10
ID Q61580 standard; DNA; 27 BP.
AC Q61580;
DE Neisseria 16S rRNA (750-776) comparable DNA.
DE Neisseria gonorrhoeae; 16S rRNA; identification;
KW Neisseria meningitidis; Neisseria lactamica;
KW strand displacement assay; ss.
OS Neisseria sp.

PN WO9406937-A.
PD 31-MAR-1994.
PF 15-SEP-1993; U08707.
PR 15-SEP-1992; US-945156.
PA (BOEF) BOEHRINGER MANNHEIM CORP.
PI Balgobin N, Buck H, Eadie S;
PI WPI: 94-118485/14.
DR New complex for determining a target nucleotide sequence -
PT comprising a probe having an initial binding region and a target
PT binding region and a hybridised labelled sequence
PS Example 1; Page 5-6; 59pp; English.
CC Studies were undertaken to prepare appropriate materials for
CC identification of N. gonorrhoeae. The 16S rRNA gene for a specific
CC strain of N. gonorrhoeae is known (Rossau et al., Nucl. Acids Res.
CC 16: 6227 (1990)). The comparable DNA for 16S rRNA of specific
CC strains of N. meningitidis and N. lactamica was sequenced
CC Comparison showed that nucleic acids 1262-1281 was non-homologous
CC (see Q61576-78). Three other regions were identified with a single
CC base pair difference (see Q61579-81).
SQ Sequence 27 BP; 6 A; 9 C; 6 G; 6 T;
Query Match 52.2%; Score 12; DB 10; Length 27;
Best Local Similarity 87.5%; Pred. No. 5.70e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 6 gcaagtcacagggtttc 21
||||| ||||| |||||
Cp 23 gcaaggacagggtttc 8
RESULT 11
ID Q61782 standard; DNA; 28 BP.
AC Q61782;
DT 31-OCT-1994 (first entry)
DE Fragment of 16S rRNA of Neisseria gonorrhoeae.
KW Neisseria gonorrhoeae; ribosome; ribosomal RNA; ribonucleic acid;
KW rRNA; Neisseria meningitidis; Neisseria lactamica; probe;
KW detection; identification; ss.
OS Neisseria gonorrhoeae.
PN WO9406817-A.
PD 31-MAR-1994.
PF 15-SEP-1993; U08702.
PR 15-SEP-1992; US-945159.
PA (BOEF) BOEHRINGER MANNHEIM CORP.
PI Chmelo R, Foltz L;
PI WPI: 94-118390/14.
DR Isolated nucleotide sequences form Neisseria gonorrhoeae - used
PT as probes in assays for specific detection of Neisseria
PT gonorrhoeae in samples
PS Example 1; Page 6; 56pp; English.
CC The sequence corresponds to bases 750-776 of the 16S rRNA of
CC Neisseria gonorrhoeae and can be used as a probe for the specific
CC identification of Neisseria gonorrhoeae. It differs from the
CC corresponding sequences from N. meningitidis and N. lactamica by
CC having adenine at position 763.
SQ Sequence 28 BP; 6 A; 9 C; 6 G; 7 T;
Query Match 52.2%; Score 12; DB 10; Length 28;
Best Local Similarity 87.5%; Pred. No. 5.70e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 7 gcaagtcacagggtttc 22
||||| ||||| |||||
Cp 23 gcaaggacagggtttc 8
RESULT 12
ID T47853 standard; DNA; 33 BP.
AC T47853;
DT 21-MAY-1997 (first entry)
DE PCR primer used to clone TEV viral proteinase site.
KW PCR; polymerase chain reaction; amplification; construction;
KW exonuclease; directional cloning; join; fuse; site-specific mutation;

KW vector; recombinant; ss.
 OS Synthetic.
 PN US5580759-A.
 PD 03-DEC-1996.
 PF 03-FEB-1994; 192300.
 PR 03-FEB-1994; US-192300.
 PI (TEXA) UNIV TEXAS SYSTEM.
 PA Capra JD, Tucker PW, Yang Y;
 DR WPI: 97-033570/03.
 PT Construction of recombinant DNA mols. - using exo:nuclease-based
 PT methods
 PS Example 6; Column 19-20; 28pp; English.
 CS T47853 and T47854 are PCR primers used in an example for directional
 CC cloning of a TEV viral proteinase site into a single restriction
 CC site. The example uses an exonuclease-based method for joining and/or
 CC constructing two or more DNA molecules. DNA fragments containing
 CC ends complementary to those of a vector were generated by PCR. The 3'
 CC ends of these molecules and vector were then reprocessed by
 CC exonuclease activity and annealed in an orientation-determinant
 CC manner via their complementary single-stranded regions. The
 CC resultant recombinant DNA may be transformed directly into bacteria
 CC without a further ligase-dependent reaction. Using this approach,
 CC recombinant DNA molecules are constructed rapidly, efficiently and
 CC directionally. This method can effectively replace conventional PCR
 CC cloning, PCR SOBing, DNA subcloning and site-directed mutagenesis
 CC protocols.
 SQ Sequence 33 BP; 8 A; 9 C; 7 G; 9 T;
 Query Match 52.2%; Score 12; DB 29; Length 33;
 * Best Local Similarity 87.5%; Pred. NO. 5.70e+02;
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Db 8 ctgaaacacctgtactt 23
 Qy 6 ctgaaacacctgtcctt 21
 RESULT 13
 ID T25660 standard; cDNA to mRNA; 35 BP.
 AC T25660;
 DT 14-DEC-1996 (first entry)
 DE Human gene signature HUMG07855.
 KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
 KW human; cloning; mapping; non-biased library; diagnosis; detection;
 KW cell typing; abnormal cell function; ss.
 OS Homo sapiens.
 PN WO9514772-A1.
 PD 01-JUN-1995.
 PF 11-NOV-1994; J01916.
 PR 12-NOV-1993; JP-355504.
 PA (MATS/) MATSUBARA K.
 PI (OKUB/) OKUBO K.
 PI Matsubara K, Okubo K;
 DR WPI: 95-206931/27.
 PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
 PT for diagnosis of abnormal cell function, by preparing cDNA that
 PT reflects relative abundance of corresp. mRNA in specific human
 PT tissues
 PS Claim 1; Page 1300; 2245pp; Japanese.
 CC A single-stranded DNA (or its complementary strand or the corresp.
 CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
 CC given in T19001-T26837 and which is able to hybridise to part of
 CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
 CC sequences were obtained from 3'-directed cDNA libraries prepared
 CC from various human tissues; synthesis of cDNA was initiated from the
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
 CC untranslated sequence is unique to a particular mRNA species, almost
 CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
 CC is constructed so as to reflect accurately the relative abundance of
 CC different mRNAs in the particular tissue from which it was derived.
 CC The appearance frequency of a given GS in a cDNA library can be
 CC determined (esp. using primers and probes derived from the GS
 CC sequences) as a means of diagnosing abnormal cell function or for

CC recognising different cell types.
 SQ Sequence 35 BP; 10 A; 11 C; 7 G; 7 T;
 Query Match 52.2%; Score 12; DB 23; Length 35;
 Best Local Similarity 100.0%; Pred. NO. 5.70e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 19 tgaacacctgtc 30
 Qy 7 tgaacacctgtc 18
 RESULT 14
 ID T52876 standard; RNA; 36 BP.
 AC T52876;
 DT 14-APR-1997 (first entry)
 DE Human ICAM hammerhead ribozyme sequence (nt. position 2659).
 KW Enzymatic nucleic acid; ribozyme; trans cleavage; inhibition;
 KW gene expression; downregulation; interleukin-5; IL-5; ICAM-1;
 KW intercellular adhesion molecule; rel A; tumour necrosis factor;
 KW TNF-alpha; respiratory syncytial virus; RSV; bcr-abl; oncogene;
 KW translocation; chronic myelogenous leukaemia; CML; cancer;
 KW Philadelphia chromosome; inflammation; autoimmune disease;
 KW atherosclerosis; myocardial infarction; stroke; restenosis;
 KW transplant rejection; rheumatoid arthritis; psoriasis;
 KW myocardial ischaemia; Kawasaki disease; septic shock; HIV;
 KW human immunodeficiency virus; acquired immune deficiency syndrome;
 KW AIDS; ss.
 OS Synthetic.
 PN WO9523225-A2.
 PD 31-AUG-1995.
 PF 23-FEB-1995; IB0156.
 PR 23-FEB-1994; US-203109.
 PR 29-MAR-1994; US-218934.
 PR 04-APR-1994; US-222795.
 PR 07-APR-1994; US-224483.
 PR 15-APR-1994; US-228041.
 PR 15-APR-1994; US-227958.
 PR 18-MAY-1994; US-245736.
 PR 06-JUL-1994; US-271280.
 PR 15-AUG-1994; US-291932.
 PR 16-AUG-1994; US-291433.
 PR 17-AUG-1994; US-292620.
 PR 19-AUG-1994; US-293520.
 PR 02-SEP-1994; US-300000.
 PR 08-SEP-1994; US-303039.
 PR 23-SEP-1994; US-311749.
 PR 23-SEP-1994; US-311486.
 PR 28-SEP-1994; US-314397.
 PR 03-OCT-1994; US-316771.
 PR 07-OCT-1994; US-319492.
 PR 11-OCT-1994; US-321993.
 PR 04-NOV-1994; US-334847.
 PR 10-NOV-1994; US-337608.
 PR 28-NOV-1994; US-345516.
 PR 16-DEC-1994; US-357577.
 PR 23-DEC-1994; US-363233.
 PR 30-JAN-1995; US-380734.
 PA (RIBO-) RIBOZYME PHARM INC.
 PI Stinchcomb DT, Chowrira B, Drenzo A, Draper KG, Dudycz LW;
 PI Grimm S, Karpelsky A, Kisich K, Matulic-Adamic J;
 PI McSwiggen JA, Modak A, Favco P, Beigelman L, Sullivan SM;
 PI Sweedler D, Thompson JD, Tracz D, Usman N, Wincott FE;
 PI Woolf T;
 DR WPI: 95-351090/45.
 PT Ribozymes having modified bases and methods for producing them
 PT for use in inhibiting disease related genes
 PS Claim 9; Page 188; 407pp; English.
 CC The present sequence is that of a claimed enzymatic nucleic acid
 CC (i.e. a ribozyme) which cleaves ICAM-1 mRNA at the nucleotide base
 CC position indicated in the DE line.
 CC Regions of the mRNA that do not form secondary folding
 CC structures and that contain potential hammerhead and hairpin

CC ribozyme cleavage sites were identified by computer analysis.
 CC Ribozymes directed against these mRNA sequences were designed and
 CC synthesised with modifications that improve their nuclease
 CC resistance. The ribozymes cleave the ICAM-1 target sequences and
 CC thereby inhibit ICAM-1 expression, making them useful for reducing
 CC transplant rejection and alleviating symptoms in patients with
 CC rheumatoid arthritis, asthma and other inflammatory disorders.
 SQ Sequence 36 BP; 10 A; 10 C; 11 G; 5 U;

Query Match 52.2%; Score 12; DB 27; Length 36;
 Best Local Similarity 85.7%; Pred. No. 5.70e+02;
 Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 22 agccgaaccccg 35
 |||||
 QY 3 aggcgaaccccg 16

RESULT 15

ID Q69355 standard; DNA; 46 BP.
 AC Q69355;
 DT 23-FEB-1995 (first entry)
 DE Human CYPXIX aromatase P-450 gene, target region.
 KW DNA protein-binding assay; test sequence; screening sequence;
 KW promoter; target; TATA box; Herpes Simplex Virus; HSV;
 KW origin of replication; UL9; transcription factor; TfIID: ds.
 OS Synthetic.
 PN WO9414980-A.
 PD 07-JUL-1994.
 PF 20-DEC-1993; U12388.
 PR 23-DEC-1992; US-996783.
 PR 17-SEP-1993; US-123936.
 PA (GENE-) GENELABS TECHNOLOGIES INC.
 PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
 DI WPI; 94-234711/28.
 DI Sequence-directed DNA-binding molecules - useful in
 PT pharmaceuticals and as molecular reagents
 P3 Claim 28; Page 264; 587pp; English.
 CC A DNA protein-binding assay is provided, useful for screening
 CC libraries of synthetic or biological cpds. for their ability
 CC to bind DNA test sequences. The assay is versatile in that any
 CC number of test sequences can be tested by placing the test sequence
 CC adjacent to a defined protein-binding screening sequence. Binding
 CC of mols. to these test sequences changes the binding characteristics
 CC of the protein mol. to its cognate binding sequence. When such a mol.
 CC binds the test sequence, the equilibrium of the DNA:protein complexes
 CC is disturbed, generating changes in the concentration of free DNA probe.
 CC One application of this method is to eucaryotic general transcription
 CC factors (e.g. TfIID) where the target region is typically selected
 CC from DNA sequences adjacent to the binding site for the eucaryotic
 CC transcription factor. Numerous exemplary test sequences are given:
 CC the sequences in Q69251-731 and Q69850 correspond to promoter targets
 CC (typically, TATA box-contg. sites) for human genes and the sequences in
 CC Q69732-849 correspond to promoter targets for viral genes. The test
 CC sequences may also be randomly generated. DNA:protein interaction may
 CC be used for screening purposes, e.g. the Herpes Simplex Virus (HSV)
 CC origin of replication and UL9 (see Q69851-52, Q69865 and Q69891).
 SQ Sequence 46 BP; 19 A; 9 C; 10 G; 8 T;

Query Match 52.2%; Score 12; DB 12; Length 46;
 Best Local Similarity 100.0%; Pred. No. 5.70e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 34 aaggacagggtt 45
 |||||
 Cp 21 aaggacagggtt 10

Search completed: Fri Apr 2 10:04:51 1999
 Job time : 225 secs.

WQELH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_rm n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Fri Apr 2 09:50:36 1999; MasPar time 77.99 Seconds
Tabular output not generated. 528.096 Million cell updates/sec

Title: >US-09-238-972-2
Description: (1-23) from US09238972.seq
Perfect Score: 23
N.A. Sequence: 1 gtagcgtgaacccctgtcctgc 23
Comp: catccgactttggacaggaacg

Scoring table: TABLE default
Gap 10

Nmatch STD : Dbase 0; Query 0

Searched: 2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 0%
Listing first 1000 summaries
Maximum DB seq length 100

Database: emb1-est56
1:em_est1 2:em_gss1 3:em_gss2 4:em_gss3
Database: genbank-est109
5:gb_est1 6:gb_est10 7:gb_est11 8:gb_est12 9:gb_est13
10:gb_est14 11:gb_est15 12:gb_est16 13:gb_est17
14:gb_est18 15:gb_est19 16:gb_est2 17:gb_est20
18:gb_est21 19:gb_est3 20:gb_est4 21:gb_est5 22:gb_est6
23:gb_est7 24:gb_est8 25:gb_est9 26:gb_gss1 27:gb_gss2
28:gb_gss3 29:gb_gss4

Statistics: Mean 6.917; Variance 1.457; scale 4.748

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description	Pred. No.
1	14	60.9	60	26	B04047 CSRL-23c11-u CSRL flow 8.26e+00	

Note: Post-processor removed 999 summaries from list due to search parameters chosen.

ALIGNMENTS

RESULT 1 B04047 60 bp DNA GSS 26-JUN-1996
LOCUS CSRL-23c11-u CSRL flow sorted Chromosome 11 specific cosmid Homo
DEFINITION sapiens genomic clone CSRL-23c11, genomic survey sequence.
ACCESSION B04047

NID g1413325
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 60)
AUTHORS Evans,G.A., Burbee,D., Davies,C., Hahner,L., Oliver,T., Gilbert,M., Jones,D., Ward,T., Gillilan,E., Schagemann,J., Probst,S., Harris,J., DeFord,J., McFarland,J., Burzinski,K., Khan,M., Kupfer,K. and Garner,H.R.
TITLE Genomic Sequence Sampled Map of Chromosome 11
JOURNAL Unpublished (1996)
COMMENT

Contact: Evans GA, Shane Probst
McDermott Center for Human Growth and Development
University of Texas Southwestern Medical Center At Dallas
5323 Harry Hines Blvd, Dallas TX 75235-8591
Tel: 214-648-1600
Fax: 214-648-1666
Email: gevas@utsw.swmed.edu, shane@mcdermott.swmed.edu
Class: cosmid ends
Seg primer: T7
High quality sequence stop: 60.
Location/Qualifiers 1..60

FEATURES source
/organism="Homo sapiens"
/note="vector: sCos-1; Human Chromosome 11 specific cosmid library prepared from flow sorted human Chromosome 11 derived from Chinese Hamster Ovary (CHO) monochromosomal somatic cell hybrid, J1"
/db_xref="taxon:9606"
/clone="CSRL-23c11"
/clone_lib="CSRL flow sorted Chromosome 11 specific cosmid"
/sex="female"
/cell_type="chimeric hamster somatic cell hybrid"
BASE COUNT 17 a 12 c 19 g 11 t 1 others
ORIGIN

Query Match 60.9%; Score 14; DB 26; Length 60;
Best Local Similarity 93.3%; Pred No. 8.26e+00;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 25 CTGANACCCCTGTCTCT 39
|||||
Qy 6 ctgaacccctgtcct 20

Search completed: Fri Apr 2 10:00:42 1999
Job time : 606 secs.

THIS PAGE BLANK (uspto)

Result No.	Score	Query		Length	DB	ID	Description	Pred. No.
		Match	%					
C 1	23	100.0	1989	29	MUSCAT2A	Mus musculus cationic	1.60e-03	
C 2	23	100.0	1989	29	MUSCAAT2X	Mouse cationic amino a	1.60e-03	
C 3	23	100.0	2397	22	I13751	Sequence 5 from patent	1.60e-03	
C 4	23	100.0	2723	29	MUSTEAB	Mouse T cell early act	1.60e-03	
C 5	23	100.0	3698	29	MUSMEMPROT	Mouse membrane protein	1.60e-03	
C 6	17	73.9	523	22	E06159	DNA fragment of mercur	1.33e+01	
C 7	17	73.9	5001	16	TFMERORRG	T.ferrooxidans mercur an	1.33e+01	
C 8	17	73.9	144868	28	H808116	Human DNA sequence fro	1.33e+01	
C 9	17	73.9	148640	21	AF056116	Fugu rubripes serine/t	1.33e+01	
C 10	17	73.9	177142	27	AC004582	Human Chromosome 11p14	1.33e+01	
C 11	17	73.9	200349	28	HSU85195	Homo sapiens BAC129.c	1.33e+01	
C 12	17	73.9	250529	27	HUAE000058	Homo sapiens T-cell re	1.33e+01	
C 13	16	69.6	366	31	GU05370	human STS WT-7508	5.26e+01	

AGVISAVMAFLFDLKALVDMMSIGTILMAYSLVAACVLMRLRYQGLCYEQPKYTPKET
LESCITNATKSEQVMTLMQOGFSRLTFLSPASLTPROQSASLVFLVGLFALILGLS
ILTYGVQAIARLEAWSALLALFLVLCRAVILTWRPONQOKVAFMVPFLPFLPAF
SILVNIYLMQLSADTWIRSEIWMALGFLIYFAYGIRHSLGNPRDEDEDDAFSENI
NVATEKSVQANDHQRNLSLFFILHKTSEC"

BASE COUNT 434 a 508 c 494 g 553 t
ORIGIN
Query Match 100.0%; Score 23; DB 29; Length 1989;
Best Local Similarity 100.0%; Pred. No. 1.60e-03;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1413 GCAAGGACAGGGTTTCAGCCTAC 1435
|||||
Cp 23 gcaaggacaggggttcagcctac 1

RESULT 2 MUSCAAT2X 1989 bp mRNA ROD 02-MAY-1994
LOCUS Mouse cationic amino acid transporter-2 mRNA, complete cds.
DEFINITION L11600
ACCESSION g293314
NID
KEYWORDS amino acid transporter; cationic amino acid transporter-2.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
Murinae; Mus.

REFERENCE 1 (sites)
MacLeod,C.L., Finley,K., Kakuda,D., Kozak,C.A. and Wilkinson,M.F.
Activated T-cells express a novel gene on chromosome 8 that is
closely related to the murine exotropic retroviral receptor
Mol. Cell. Biol. 10, 3663-3674 (1990)
90287157
REFERENCE 2 (bases 1 to 1989)
Closs,E.I., Albritton,L.M., Kim,J.W. and Cunningham,J.M.
Identification of a low affinity, high capacity transporter of
cationic amino acids in mouse liver
J. Biol. Chem. 268, 7538-7544 (1993)
93216705

FEATURES
source Location/Qualifiers
1..1989 /organism="Mus musculus"
/db_xref="taxon:10090"
/dev_stage="adult"
/tissue_type="liver"
/tissue_lib="lambda GT10"
10..1983 /standard_name="MCAT-2"
/note="putative"
/codon_start=1
/product="cationic amino acid transporter-2"
/db_xref="PDB:g293315"

CDS

/translation="MIPCAVLTFAFLIRKIVLDSLEDSKLCRLTVDLIAGLV
GSTLGVAVLAGEKADSGPSIVVFLIAALASVMAGLCYAEFGARPKTGSAYL
TYVTGELWAFITGNLILSYVIGTSVAVRANGTFDELLNKOIGOFFKTYKMYTG
LAEPDFFAVCLVLLGLLSFGVKESAVNKFETAINILVLLFVMVAGFVGNVAV
KTSEELKLSASAREPPSENGSTIYGAGGEMPGFTGTLAGAANTCFYAFVGFDCIAT
TGEEVNPKAIPIGIVTSILVCFMAYFGVSAALTLMPYILLDEKSPLPVAFETVRW
GKAYVAVGSCALSTLGSMPFLPRILFAMARDGLLFRFLARVSKRQSPVAATMT
AGVISAVMAFLFDLKALVDMMSIGTILMAYSLVAACVLMRLRYQGLCYEQPKYTPKET
LESCITNATKSEQVMTLMQOGFSRLTFLSPASLTPROQSASLVFLVGLFALILGLS
ILTYGVQAIARLEAWSALLALFLVLCRAVILTWRPONQOKVAFMVPFLPFLPAF
SILVNIYLMQLSADTWIRSEIWMALGFLIYFAYGIRHSLGNPRDEDEDDAFSENI
NVATEKSVQANDHQRNLSLFFILHKTSEC"

BASE COUNT 434 a 508 c 493 g 554 t
ORIGIN
Query Match 100.0%; Score 23; DB 29; Length 1989;
Best Local Similarity 100.0%; Pred. No. 1.60e-03;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1413 GCAAGGACAGGGTTTCAGCCTAC 1435

|||||
Cp 23 gcaaggacaggggttcagcctac 1

RESULT 3 I13751 2397 bp DNA PAT 08-AUG-1995
LOCUS Sequence 5 from patent US 5440017.
DEFINITION I13751
ACCESSION g996817
NID
KEYWORDS
SOURCE Unknown.
ORGANISM Unclonified.

REFERENCE 1 (bases 1 to 2397)
AUTHORS MacLeod,C.L.
TITLE T-cell lymphoma cDNA clones
JOURNAL Patent: US 5440017-A 5 08-AUG-1995;
FEATURES Location/Qualifiers
1..2397 /organism="unknown"
BASE COUNT 548 a 571 c 587 g 691 t
ORIGIN

Query Match 100.0%; Score 23; DB 22; Length 2397;
Best Local Similarity 100.0%; Pred. No. 1.60e-03;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1201 GCAAGGACAGGGTTTCAGCCTAC 1223
|||||
Cp 23 gcaaggacaggggttcagcctac 1

RESULT 4 MUSTEAB 2723 bp mRNA ROD 08-SEP-1995
LOCUS Mouse T cell early activation protein (tea) mRNA, complete cds.
DEFINITION M62838 M32485
ACCESSION g976160
NID
KEYWORDS T-cell early activation protein; cell surface glycoprotein.
SOURCE Mouse (strain AKR/J) CDNA to mRNA.
ORGANISM Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
Murinae; Mus.

REFERENCE 1 (bases 327 to 2723)
AUTHORS MacLeod,C.L., Finley,K., Kakuda,D., Kozak,C.A. and Wilkinson,M.F.
TITLE Activated T cells express a novel gene on chromosome 8 that is
closely related to the murine ecotropic retroviral receptor
Mol. Cell. Biol. 10 (7), 3663-3674 (1990)
90287157

REFERENCE 2 (bases 1 to 2723)
AUTHORS Reizer,J., Finley,K., Kakuda,D., MacLeod,C.L., Reizer,A. and
Saiter,M.H. Jr.
TITLE Mammalian integral membrane receptors are homologous to
facilitators and antiporters of yeast, fungi, and eubacteria
protein Sci. 2 (1), 20-30 (1993)
93184721

REFERENCE 3 (bases 1 to 2723)
AUTHORS MacLeod,C.L.
TITLE Direct Submission

JOURNAL Submitted (01-DEC-1990) Carol L. MacLeod, Cancer Genetics Program,
UCSD Cancer Center, La Jolla, CA 92093, USA
On Sep 8, 1995 this sequence version replaced gi:476724.
FEATURES Location/Qualifiers
1..2723 /organism="Mus musculus"
/strain="AKR/J"
/db_xref="taxon:10090"
/clone="20.5"
/cell_line="SL12.4.10"
/tissue_type="T cell lymphoma"
/chromosome="8"
121..2097 /gene="tea"

source

gene


```
CDS      121..2097
/gene="tea"
/codon_start=1
/product="T-cell early activation protein"
/db_xref="PID:g476725"
/translation="MIPRAVLTFARCLIRKIVTLDSLEDSKLCRLTITVDLIALGV
TSTIGAGVYVLAVGAKADSPISVSEFLAALASVMAGLCYAEFGARVPTGSAYLY
GYTVGELMAFITGNLILSVIGTSSVARWSGTFDELLNKQIGQFKTYKNNYTG
LAEPDFKAVLITGNLLAGLSFGKESAWNKFTAINILVLFVMVAGVKNVANN
KISEFKNTISASAREPPSENGTSIYGAGGFMPTGTLAGAATCKYAFVGFDCIAT
TGSEVRNPQRAIGIVTSLVCFMAYGSAULTLMPYLLDEKSPLPVAFYVRW
GPKYVVAAGSLCALSTSLGSIFFMPRIYVMAEDGLLKFCLAQINSKTKTPVIALT
SSGAAVMAFLFDKALYDMSIGTLMAYSLVAACVLLILRYQGLCYEQPKYPEKE
TLESCNATKESQVMTLOGGFSKTLFSPALPQRSASLVSLVGLFLAFILGL
SILTYGQVAILAEWSLALLFLVLCRAVILTIWRQPNQOKVAFWPFPLPFA
FSILVNIYLMVQLSADTWIRFSIMWALGFLIYFAYGIRHSLEGNPRDEDEDAFSEN
INVATEKSVMAQNDHQHNLSPFILHEKTS"
BASE COUNT      613 a      664 g      778 t
ORIGIN
Query Match      100.0%; Score 23; DB 29; Length 2723;
Best Local Similarity 100.0%; Pred. No. 1.60e-03;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1527 GCAAGGACAGGGTTTCAGCCTAC 1549
|||||
Cp 23 gcaaggacagggtttcagcctac 1

RESULT 5
LOCUS      MUSEMPROT 3698 bp mRNA ROD 01-AUG-1994
DEFINITION Mouse membrane protein (Tea) mRNA, complete cds.
ACCESSION L29006
NID      g456102
KEYWORDS membrane protein.
SOURCE      Mus musculus
ORGANISM      Mus musculus
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
Murinae; Mus.
REFERENCE 1 (bases 1 to 3698)
AUTHORS      Kavanagh,M.P., Wang,H., Zhang,Z., Zhang,W., Wu,Y.N., Dechant,E.,
North,R.A. and Kabat,D.
TITLE      Control of cationic amino acid transport and retroviral receptor
functions in a membrane protein family
JOURNAL      J. Biol. Chem. 269 (22), 15445-15450 (1994)
MEDLINE      94253120
FEATURES
source      Location/Qualifiers
1..3698
/organism="Mus musculus"
/db_xref="taxon:10090"
/tissue_type="liver"
/tissue_lib="lambda gtl1; Clontech"
533..2506
/gene="Tea"
533..2506
/gene="Tea"
/feature="T lymphocyte early-activation gene"
/codon_start=1
/product="membrane protein"
/db_xref="PID:g517493"
/translation="MIPRAVLTFARCLIRKIVTLDSLEDSKLCRLTITVDLIALGV
TSTIGAGVYVLAVGAKADSPISVSEFLAALASVMAGLCYAEFGARVPTGSAYLY
GYTVGELMAFITGNLILSVIGTSSVARWSGTFDELLNKQIGQFKTYKNNYTG
LAEPDFKAVLITGNLLAGLSFGKESAWNKFTAINILVLFVMVAGVKNVANN
KISEFKNTISASAREPPSENGTSIYGAGGFMPTGTLAGAATCKYAFVGFDCIAT
TGSEVRNPQRAIGIVTSLVCFMAYGSAULTLMPYLLDEKSPLPVAFYVRW
GPKYVVAAGSLCALSTSLGSIFFMPRIYVMAEDGLLKFCLAQINSKTKTPVIALT
SSGAAVMAFLFDKALYDMSIGTLMAYSLVAACVLLILRYQGLCYEQPKYPEKE
TLESCNATKESQVMTLOGGFSKTLFSPALPQRSASLVSLVGLFLAFILGL
SILTYGQVAILAEWSLALLFLVLCRAVILTIWRQPNQOKVAFWPFPLPFA
FSILVNIYLMVQLSADTWIRFSIMWALGFLIYFAYGIRHSLEGNPRDEDEDAFSEN
INVATEKSVMAQNDHQHNLSPFILHEKTS"

gene
CDS      533..2506
/gene="Tea"
/feature="T lymphocyte early-activation gene"
/codon_start=1
/product="membrane protein"
/db_xref="PID:g517493"
/translation="MIPRAVLTFARCLIRKIVTLDSLEDSKLCRLTITVDLIALGV
TSTIGAGVYVLAVGAKADSPISVSEFLAALASVMAGLCYAEFGARVPTGSAYLY
GYTVGELMAFITGNLILSVIGTSSVARWSGTFDELLNKQIGQFKTYKNNYTG
LAEPDFKAVLITGNLLAGLSFGKESAWNKFTAINILVLFVMVAGVKNVANN
KISEFKNTISASAREPPSENGTSIYGAGGFMPTGTLAGAATCKYAFVGFDCIAT
TGSEVRNPQRAIGIVTSLVCFMAYGSAULTLMPYLLDEKSPLPVAFYVRW
GPKYVVAAGSLCALSTSLGSIFFMPRIYVMAEDGLLKFCLAQINSKTKTPVIALT
SSGAAVMAFLFDKALYDMSIGTLMAYSLVAACVLLILRYQGLCYEQPKYPEKE
TLESCNATKESQVMTLOGGFSKTLFSPALPQRSASLVSLVGLFLAFILGL
SILTYGQVAILAEWSLALLFLVLCRAVILTIWRQPNQOKVAFWPFPLPFA
FSILVNIYLMVQLSADTWIRFSIMWALGFLIYFAYGIRHSLEGNPRDEDEDAFSEN
INVATEKSVMAQNDHQHNLSPFILHEKTS"

LOCUS      TFERRCG 5001 bp DNA BCT 30-JUN-1993
DEFINITION T.ferrooxidans merR and merC genes.
ACCESSION X57326 S79623
NID      g48150
KEYWORDS merC gene; mercuric ion resistance; merR gene.
SOURCE      Thiobacillus ferrooxidans.
ORGANISM      Eubacteria; Proteobacteria; gamma subdivision.
```

```
BASE COUNT      873 a      897 c      872 g      1056 t
ORIGIN
Query Match      100.0%; Score 23; DB 29; Length 3698;
Best Local Similarity 100.0%; Pred. No. 1.60e-03;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1936 GCAAGGACAGGGTTTCAGCCTAC 1958
|||||
Cp 23 gcaaggacagggtttcagcctac 1

RESULT 6
LOCUS      E06159 523 bp DNA PAT 26-NOV-1996
DEFINITION DNA fragment of mercury-resistant gene in pTMR15.
ACCESSION E06159
NID      g2174346
KEYWORDS JP 1994000083-A/3.
SOURCE      Thiobacillus ferrooxidans.
ORGANISM      Thiobacillus ferrooxidans.
Eubacteria; Proteobacteria; gamma subdivision.
REFERENCE 1 (bases 1 to 523)
AUTHORS      Juichi,S., Chihiro,I., Satoyuki,I., Masahiko,N., Tomonobu,K. and
Kazuyuki,S.
TITLE      GENE FOR ENHANCING TRANSFORMATION, USING MERCURY RESISTANT SHUTTLE
VECTOR PLASMID FOR THE GENUS THIOBACILLUS
JOURNAL      Patent: JP 1994000083-A 3 11-JAN-1994;
COMMENT      DOWA MINING CO LTD, AKITA PREF GOV
OS      Thiobacillus ferrooxidans
PN      JP 1994000083-A/3
PD      11-JAN-1994
PF      17-JAN-1991 JP 1991018338
PI      SHIRATORI JUICHI, INOUE CHIHIRO, TAKESHIMA SATOYUKI, PI
NUMATA MASAHIKO,
PI      KUSANO TOMONOBU, SUGAWARA KAZUYUKI
PC      C12N15/31,C12N1/21,C12N15/53,C12N15/74//C12N9/02,(C12N1/21,PC
C12R1:19);
PC      (C12N1/21,C12R1:01),(C12N15/74,C12R1:01,C12R1:19); CC
strandedness: Single;
CC topology: Linear;
CC *source: strain=E-15;
CC *source: library=BamHI Library of genome DNA of
C12R1:19);
FH Key      Location/Qualifiers
FT RBS      101..108
FT CDS      116..523
FT
FT
FEATURES
source      Location/Qualifiers
1..523
/organism="Thiobacillus ferrooxidans"
/db_xref="taxon:920"
BASE COUNT      114 a      128 c      154 g      127 t
ORIGIN
Query Match      73.9%; Score 17; DB 22; Length 523;
Best Local Similarity 94.7%; Pred. No. 1.33e+01;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 285 AAGGACTGGGTTTCAGCCT 303
|||||
Cp 21 aaggacagggtttcagcct 3

RESULT 7
LOCUS      TFERRCG 5001 bp DNA BCT 30-JUN-1993
DEFINITION T.ferrooxidans merR and merC genes.
ACCESSION X57326 S79623
NID      g48150
KEYWORDS merC gene; mercuric ion resistance; merR gene.
SOURCE      Thiobacillus ferrooxidans.
ORGANISM      Eubacteria; Proteobacteria; gamma subdivision.
```

```

REFERENCE
AUTHORS      Kusano,T.
TITLE        Direct Submission
JOURNAL      Submitted (22-JAN-1991) T. Kusano, Biotechnology Institute, Akita
              Prefectural College of Agriculture, Oghata Minami-Akita Gun, Akita
              010 04, JAPAN
REFERENCE
AUTHORS      Inoue,C., Sugawara,K. and Kusano,T.
TITLE        The merR regulatory gene in Thioabacillus ferrooxidans is spaced
              apart from the mer structural genes
JOURNAL      Mol. Microbiol. 5 (11), 2707-2718 (1991)
MEDLINE      92140035
FEATURES
source       1. .5001
              /organism="Thiobacillus ferrooxidans"
              /strain="E-15"
              /db_xref="taxon:920"
              complement(612. .1019)
              /gene="merR2"
              complement(612. .1019)
              /gene="merR2"
              /function="regulatory gene"
              /codon_start=1
              /transl_table=11
              /db_xref="PID:g48151"
              /translation="MSRLDTGTGVLGVSGLSACMPALASIGAAGVGLGFLSRWEG
              SDLERLHFVKSAGLGFSLSEIGQLKLADGTHCREAAELASRHLASVQARLRELHRI
              EVALQKLEACNSQGNFSCPLDLSRLKTVS"
              1135. .1157
              /gene="merC2"
              1135. .1157
              /gene="merC2"
              /function="mercuric ion transport"
              /codon_start=1
              /transl_table=11
              /db_xref="PID:g48152"
              /translation="MSRLDTGTGVLGVSGLSACMPALASIGAAGVGLGFLSRWEG
              LFVHPLIPVFLVIALLANGLGFHGHQORSALGCGVPVALLFGQGTQHFLPVEMA
              RILFTGLVIMVIAIWMFRPAHAKCPVPSPGDNQSHC"
              1371. .2392
              /note="ORF-1"
              /codon_start=1
              /transl_table=11
              /db_xref="PID:g48153"
              /translation="MSQIFLRITGMTCAHCAHSEKALLGTHGIDSAQVSLATNOAE
              VFLQSSIPTEALLAATQAGYAKVEQDSLOQARSTQEPQPHAIIGSGGAAMAA
              LKAVRGARVTLIARTIGGTCVKVCPSKILIRAAHIAHLRRASPDAGTAAQEPV
              IRRDLAQOQGRVDLYAKYEGILAQTPAISVLRGATFONOTLSVVLADGVHE
              LRFDLCLTAGSPAIPPLGLADTPYMTSTALASETIPERLVIGASVVAVELAQH
              FARLGSKVTLARSVLLSHEDPAIGEATISAFRAPIAHVLEHTOPNRADYNDGQFLIT
              TNGHLRADRLLIATGRAPNTHSLNLAANGVFNAGHAIVIDXTWRTSNPYIRDLKRR
              LGGGRKLLANVRNTGSIHIEEVAQLQEGVVHTARFSAFYTDKQAPRYVAHG
              EPRFAFLATASFFQFARNGP"
              complement(2603. .3019)
              /note="ORF-2"
              /codon_start=1
              /transl_table=11
              /db_xref="PID:g48154"
              /translation="MSKGSRRSEGAIAARRLKEGRGQSEKSPKWLTVRDVPSRGLS
              VRIKRGKTRVHLLSQLELSYFLMLDDIRACGVTDIREQFPLTIETTLTETDIRCA
              GAHRLVDDGSCVCDLNAHRKVKQAMRIIRRTASGQ"
              complement(3146. .3832)
              /note="ORF-3"
              /codon_start=1
              /transl_table=11
              /db_xref="PID:g48155"
              /translation="MSQIFLRITGMTCEHCAHSEKALLGTHGVDSAQVSLATNOAE
              VFLQSSIPTEALLAATQAGYAKVEQDSLOQARSTQEPQPHAIIGSGGAAMAA
              LKAVRGARVTLIERTIGGTCVNVGCVPSKILIRAAHIAHLRRASPDAGTAAQEPV
              IRRDLAQOQGRVDLYAKYEGILAQTPAISVLRGATFONOTLSVVLADGVHE
              LRFDRCLTIAT"
              complement(3846. .4268)
              /gene="merC1"

CDS
complement(3846. .4268)
/gene="merC1"
/function="mercuric ion transport"
/codon_start=1
/transl_table=11
/db_xref="PID:g48156"
/translation="MSRLDTGTGVLGVSGLSACMPALASIGAAGVGLGFLSRWEG
LFVHPLIPVFLVIALLANGLGFHGHQORSALGCGVPVALLFGQGTQHFLPVEMA
RILFTGLVIMVIAIWMFRPAHAKCPVPSPGDNQSHC"
4384. .4791
/gene="merR1"
4384. .4791
/gene="merR1"
/function="regulatory gene"
/codon_start=1
/transl_table=11
/db_xref="PID:g48157"
/translation="MKLLTIGALAAASAGVHVTVFYORKGLLPEPDRLPISIRRYGO
SDLERLHFVKSAGLGFSLSEIGQLKLADGTHCREAAELASRHLASVQARLRELHRI
EHALQKLEACNSQGNFSCPLDLSRLKTVS"
BASE COUNT 1102 a 1451 c 1368 g 1080 t
ORIGIN
Query Match 73.9%; Score 17; DB 16; Length 5001;
Best Local Similarity 94.7%; Pred. No. 1.33e+01;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 832 AGCGTGAACCCAGTCCTT 850
|||||
QY 3 aggcgtgaacacctgtcctt 21

RESULT 8
LOCUS      HS80119 144868 bp DNA PRI 13-OCT-1998
DEFINITION Human DNA sequence from clone 80119 on chromosome 6p21.31-22.2
              Contains genes and pseudogenes for olfactory receptor-like
              proteins, STS, GSS, complete sequence.
ACCESSION  AL022727
NID        93093312
KEYWORDS   HTG.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominoidea; Homo.
            1 (bases 1 to 144868)
            Evans,K.
            Direct Submission
            Submitted (03-SEP-1998) E-mail enquiries: humquery@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
            This sequence is the entire insert of clone 80119. This sequence
            has been finished according to sequence map criteria as follows. An
            attempt is made to resolve all sequencing problems, such as
            compressions and repeats, but not necessarily within known
            annotated human repeat sequence elements (e.g. Alu). Where the
            sequence is ambiguous, there is an annotation using the 'unsure'
            feature key.
            80119 is from the library RPC11 constructed at the Roswell Park
            Cancer Institute by the group of Pieter de Jong. For further
            details see http://bacpac.med.buffalo.edu/ VECTOR: pcypac2
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above. This sequence was
            generated from part of bacterial clone contigs of human chromosome
            6, constructed in collaboration by the Sanger Centre chromosome 6
            mapping group and Armin Volz & Andreas Ziegler. Further information
            can be found at http://www.sanger.ac.uk/HGP/Chr6/.
            Location/Qualifiers
            1. .144868
            /organism="Homo sapiens"
            /db_xref="taxon:9606"

```



```
repeat_region 40921..41266
/notes="MLR1D repeat: matches 1..375 of consensus"
repeat_region 41197..41362
/notes="MLR1C repeat: matches 289..464 of consensus"
repeat_region complement(41364..41596)
/notes="MER42c repeat: matches 621..397 of consensus"
repeat_region 41598..41838
/notes="AluX repeat: matches 44..298 of consensus"
repeat_region complement(42277..42376)
/notes="MLR1D repeat: matches 333..234 of consensus"
repeat_region complement(42367..42546)
/notes="MLR1D repeat: matches 210..29 of consensus"
repeat_region 42598..44207
/notes="L1 repeat: matches 3776..5390 of consensus"
Note: remainder of annotations omitted.

Query Match 73.9%; Score 17; DB 28; Length 144868;
Best Local Similarity 90.5%; Pred. No. 1.33e+01;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 3868 AGGCTGAAGTCTCTCTTGC 3888
||||||| |||||||||
Qy 3 aggcgaacacctgtccttcg 23

RESULT 9 AF056116 148640 bp DNA VRT 05-SEP-1998
LOCUS Fugu rubripes serine/threonine kinase receptor type1, All-1 related
DEFINITION protein (ALR), fugu hedgehog (fhh), ixaros-like, wnt1, wnt10b,
ARF3, erbB3, PAS1, and L41 ribosomal protein genes, complete cds:
LRP1 gene, partial cds; and unknown genes.
ACCESSION AF056116 U82608
NID g3170534
KEYWORDS Fugu rubripes.
SOURCE Fugu rubripes.
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Actinopterygii;
Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
Tetraodontiformes; Tetraodontidae; Tetraodontidae; Fugu.
REFERENCE 1 (bases 1 to 148640)
AUTHORS Gelliner,K. and Brenner,S.
TITLE Analysis of 148 kb of genomic DNA around the wnt1 locus of Fugu
rubripes
JOURNAL Unpublished (1998)
REFERENCE 2 (bases 1 to 148640)
AUTHORS Gelliner,K. and Brenner,S.
TITLE Direct Submission
JOURNAL Submitted (30-MAR-1998) Molecular Sciences Institute, 2168 Shattuck
Ave, Berkeley, CA 94704, USA
COMMENT On Jun 2, 1998 this sequence version replaced gi:1814290.
Prediction of protein sequences are solely based on homology search
analysis and XGRAL analysis of the genomic DNA sequence.
FEATURES
source
1..148640
/organism="Fugu rubripes"
/db_xref="taxon:3103"
/clone="29P14; 3A8; 138E3; 186G5; 178O23"
/notes="sequenced from overlapping cosmids of the HGMP Fugu
cosmid library"
join(449..345,702..943,1038..1281,1393..1623,1735..1902,
2006..2162,2229..2353,2432..2562,2649..2774)
/codon_start=1
/product="serine/threonine kinase receptor type1"
/db_xref="PDB:3G170533"
/translation="MLQWRHGRSFLLLGLFTYTSWHAASSVCLALRCMCTNCKET
GYECETDGACASTYVERKEQHIRICNRDSLVPVGPFCYCLSAEGLNTHCCVEDY
CNSIDLPVCKEGEWASSGVPVELVAVIAGVFLVLLVMVFLFYOHAYSHRO
RLEVEDPCDHYLPKDKTLQDLIYDMSTSGSGSLPLFVQRTVARTIVLOEIKGR
FCEVNRGKRGDVAVKLFSSREERSWPERAIYOTIMLRHENILGFTAADNKNGTW
TOLWLSYDHKNGSLFDYLNRYSVTIEGMKRLASAGSLAHLEHILGTOCKPGIAH
RDLKSKNLLVHKGSCAIADLGLAVRHESITDITADNPQRTKRYMAPELDITIN
MKHDFSFCADIYALGLVWEIARACNAGGITHIEYOLPYDLVPSDDPSTEEMRKVYCD
```

gene

CDS

```
OKLRPNVPNWQSYEALRYMGKIMRECYANGAARLTALRIKKTLSQLSVEEDIKM"
complement(7874..29042)
/genes="ALR"
complement(join(7874..7966,8133..8241,8314..8367,
8531..8816,8931..9061,9157..9293,9381..10242,10375..10740,
10871..10939,11300..11435,11512..11636,11724..11896,
11927..12058,12137..12284,12369..12533,12619..12762,
12819..15788,15881..16113,16195..16261,16348..16429,
16726..16849,16931..18699,18902..19041,19230..19421,
19507..21090,21210..21260,21606..21679,21759..21997,
22095..22164,22287..22442,22670..22982,23076..23179,
23295..23417,23512..23688,23787..24006,24177..24300,
24454..24618,24844..25025,25108..25212,25295..25408,
25539..25646,25735..26090,26209..26301,26745..27040,
27126..27304,27408..27522,27613..27781,27884..28046,
28160..28269,28364..28566,28677..28779,28994..29042))
/genes="ALR"
/codon_start=1
/product="All-1 related protein"
/db_xref="PID:g3540281"
/translation="MDEKSNCEENDESEPTADDNASSKOLEEDSKTCTAAEDVSGSTV
ASSSTHIESVQVQALCNCEVMSLHGQRELRYGFPESEWRTLPQSTPLPQNDLDS
IGFVLPCLAAELDDSGCWVHHMCAYSGVQKQENDKLDKDVAKVIGSGIPRLCEHC
KRLGATIQCHAECCSRFYHFCPSAASGSGQMKOLLCLPEHIDAKLEACCAVC
DSAGELSDLLFCGCGHYHAACLEIGATPIQACWQCECKVCOTCKRKEDESKMLV
CDACKGYHTFCLOPAMSLPTDPHKCRRCVCTOCGARGLELPGSTOFNFAVCEA
COHHRNCTCSVNCNRPDGSVATLQSCSVRLVHSGCTLPKLESDKCILCLCKEQLPV
TOPHTAETITREAPDETAGRVDLIETITQDAAMTTEHMDVPEVTPRHKLAEEDQI
EASANTPTMDLGPQDKETTSVEQALNSHDVPTVNTOLGTSPLSHSIEEEDDD
DPLREERCLVIKLOEQKTPDLDLDETSLNLSHGDESSGFLSGPEPUAHLNMEF
GLESACSHADNLLTETDLSLPEPLRSDREKVRKSPGRSRMKQSSKFPGRRRP
RGGGGRGRGSRSLKAMASIDALSMSADTGVYKEEEDDTHQNTVILFNTDKF
VLLQDMCVVSGSFGSGEGQLLACQACQHYPCVNSKITKRLKRWKRLCELTVC
MGRASDPSRLLLDCDVSHTYCLDPLPHNVKPGKWKCKVCQCGNTNPGFHC
EMONNTHCGPCASLTPVCVRENEMFELLQCOYCDRWVHVAVESLYTEDEVAS
DEGFACVAPYKPVKGVKSNLSIFANISSTPEOYVRLGVWLTESGSLISLMS
PLHKRRRRSLGLTCCGEGADMDELREVEGDEGEGKEGCEKMECNMLNGLSPERAG
GEKADACADGAMADCDVYLGDDTDSKRRKRPYRGIGGFVNRQKCHTRKKEF
FAQLAGETILGQPIERTIDENIMDPAGEDEQAKRRKRLKMDMFPAYLQEA
FYGKTLIDLKARLVLPPGASCLVPSLPAPSGLRITPECESNQSIFFILESQ
KPYCEVTQSFFFAADANHVAKDIFPKQEGSQOAKDGTGLPGVQNDSEQF
FRKVLVSGDSGLGGMKPLESSGESHTLPQSALLPGLPSADMAFFGLSQSPF
LDMRDGGLFSPDGKEESPATPSTPVTSPSPPTETEGDGLSNQSLORWEXDEE
LGEISTISPLVXANTNFTLARDYPDMA SRCKQIMKIRKVSADKVPYLKAKNRA
AQRISKAQKAEQVCRPIKTEPGVRKERNLHLKIPLPAGVSASSOPSSAESFPFL
LPDSSSVFFSDGVPFSAEINTDPLAKFPSPCHSHPTTFFSHAGASPLQAS
FSGVYSPGPPGPPGASLPGFDMQPGTPTPRAAQVDYFRSLQKQCHLPQTQ
OGSOESLAPGSPHSRVAGIGESPLFSPSHSTHYGDAFRNQGMGRPRGVSPPSHSG
ISSPASTQYRADMSVPSRSTGTDLSTGSPAGMLESGLFLKAPMTRMHOCDGV
ALHPGASPSHSEGYKOSPSPFPESPLIPRQSGDNCGLPQRHPINQECRVPVS
SPQSHNSQSPHTPGHSDNGISAGSPATPRPSPEHCSQSPSRSHDAFAVOKPY
RSPVAPAPAFKNSPHTNSTLGLSGKPSAPPHFSIPTGGFQITQJNQMWQ
QLQSOAQONIGPDNYGARVTPSGTQEVYVROPDTHQTLPTGQMSDSTVQDP
ALGLSSELEKHKRQLREFLIRQOMORNSIKQEKAAAAAANAAGVSSGWT
GGEICAFODKTHRAPPYPODRYVTSAGTQVACVACVAVGLGDLIRPPEMG
PAIMDPTLRQGPSRQGMENRPFPPHWDQSTGPRFPQLOQANGIENLNPA
VNWQEGNLNHTIAGHGVPMQSGPPPTIELRHNAQLRPLRPFMRGPQPR
ARLVPOOTMSAIPISOHPSIQTGSIOTDGTATNSOMGLQGLSVLLPQPTGVSYTH
SHMGPOAASSPNVGTQSPQSVTVTRPQTTVENSEELPELGLGASADGV
EDEDLALDLPDGDGDLNLDNLTDPHLLDNLNDEDFDLAYTPELDQGDPKD
VFSQRLRVEATEAPSGSAGVVEIKVEQSKSAVHSTAGVCANOLPASSTAGN
LKIVDEGGLIPQVQRIQVDEIGEAVSALLGCTSSPKSTQENPASLSVRVGG
LSYPLPQATDLPHFPFGSDADDALEALPDVGGSHSPADLAKVSSLDGELLLIQD
LLEHEKQEQQLQSLSHOGGVAPHSALTNOQPNPQVAGIMLPHPHRRPQGMW
HPKKVPRPSHVQLQOQQLMGPLVPPPHMNOQQTMRMGOPGIHAGLHGQ
QPQSGVQPLRSLNNFFPKDIDKFTDIDIMPIAKAKVALKINRVAQDPMVYAPKH
OVSQVSLLAADAPATDAGLPSPGPKGETSDPTQSRPNPQFVQGIINDAEKH
QYEWLHTQQLQMLKFLFEQIGVHRKSKALCAQRTAKKAGFEADAELKL
VTEQSQIKQOLDQVRQKQKHTNLVYRSKQHQHQSSLLNPNHSGSPACAHMFP
KMPQGVQGGQAVMGHPTMPQAGMVRMPQGVFGVQPOLPLNLSGVVRGP
GTATPAGFPOGPMQSPARLLQERLOHRMQAKLQOQOQIMMGQPIPHAGNS
QNTLIPQTCMIGNVMAQOVNAQOQSGHOGSTQGMQIPQGVGSOTVYSLPQNL
AQPIIHAQAIAQPGIMNGQVAMSEQQRQMQLSQQGMVSGPGHPIRGFHSHTP
```


cloning end: 1..8775, pDJ1082L12.
 Further information regarding the map of this region or annotation
 of pDJ59m18 can be found at <http://gestec.svmed.edu/chromosome2.htm>.
 CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome 11p14.3
 Usher(USHIC) syndrome region mapped between STSs D11S1310 and
 1115A14.
 MARKER CONFIRMATION: STSs D11S1310 and SHGC-19182 sequence
 confirmed
 MAPPED CLONE OVERLAP: pDJ1082L12 and 6-130a9.

FEATURES

Source	Location/Qualifiers
repeat_region	1..17142 /organism="Homo sapiens" /db_xref="taxon:9606" complement(4265..4373)
repeat_region	/rpt_family="MIR" 9753..9910
repeat_region	/rpt_family="Alu" 11556..11768
repeat_region	/rpt_family="MIR" complement(14334..14421)
repeat_region	/rpt_family="MIR" complement(14522..14714)
repeat_region	/rpt_family="MIR" 16515..16610
repeat_region	/rpt_family="MLT1" 20835..21129
repeat_region	/rpt_family="Alu" complement(28123..28280)
repeat_region	/rpt_family="THE1" 34856..35128
repeat_region	/rpt_family="Alu" 35238..35842
repeat_region	/rpt_family="Alu" complement(37135..37431)
repeat_region	/rpt_family="Alu" complement(37443..37604)
repeat_region	/rpt_family="MIR" complement(39799..40053)
repeat_region	/rpt_family="Alu" 40438..40515
repeat_region	/rpt_family="MIR" 40578..40876
repeat_region	/rpt_family="MLT1" 41020..41217
repeat_region	/rpt_family="MLT1" complement(42907..43294)
repeat_region	/rpt_family="MER42" 43941..44235
repeat_region	/rpt_family="Alu" complement(44434..44545)
repeat_region	/rpt_family="MLT1" complement(51035..51177)
repeat_region	/rpt_family="MER5" 52054..52411
repeat_region	/rpt_family="THE1" 52445..53966
repeat_region	/rpt_family="MSTAR" 53983..54346
repeat_region	/rpt_family="THE1" complement(57581..57876)
repeat_region	/rpt_family="Alu" complement(62355..62493)
repeat_region	/rpt_family="MIR" 64734..64927
repeat_region	/rpt_family="MER30" 67371..67426
repeat_region	/rpt_family="MIR" complement(69651..69734)
repeat_region	/rpt_family="MIR" complement(72095..72210)
repeat_region	/rpt_family="MIR" complement(72639..72782)
repeat_region	/rpt_family="MIR" complement(118866..118996)
repeat_region	73013..73146 /rpt_family="MIR"
repeat_region	73640..73959 /rpt_family="Alu"
repeat_region	complement(74807..74954) /rpt_family="MIR"
repeat_region	75341..75433 /rpt_family="MER44C"
repeat_region	75448..75658 /rpt_family="MER44C"
repeat_region	76345..76675 /rpt_family="MIR"
repeat_region	complement(84972..86432) /rpt_family="MER7"
repeat_region	complement(86435..86629) /rpt_family="Alu"
repeat_region	complement(86632..86977) /rpt_family="THE1"
repeat_region	complement(86978..88519) /rpt_family="MSTAR"
repeat_region	complement(88552..88906) /rpt_family="THE1"
repeat_region	complement(88908..88988) /rpt_family="Alu"
repeat_region	89703..89772 /rpt_family="MLT1"
repeat_region	89833..90211 /rpt_family="THE1"
repeat_region	90212..90583 /rpt_family="MLT1"
repeat_region	91164..91399 /rpt_family="Alu"
repeat_region	91416..91647 /rpt_family="L1"
repeat_region	93141..93267 /rpt_family="MIR"
repeat_region	complement(94854..95076) /rpt_family="MIR"
repeat_region	complement(97298..97498) /rpt_family="MIR"
repeat_region	complement(100857..101045) /rpt_family="MIR"
repeat_region	102052..102354 /rpt_family="Alu"
repeat_region	complement(103360..103505) /rpt_family="MIR"
repeat_region	complement(105279..105514) /rpt_family="Alu"
repeat_region	106255..106562 /rpt_family="Alu"
repeat_region	107761..108120 /rpt_family="MER1"
repeat_region	108430..109760 /rpt_family="MER42"
repeat_region	110182..110221 /rpt_family="MER42"
repeat_region	complement(110515..110629) /rpt_family="Alu"
repeat_region	110649..110926 /rpt_family="MER42"
repeat_region	112736..112780 /rpt_family="Alu"
repeat_region	113214..113313 /rpt_family="MIR"
repeat_region	complement(114256..114358) /rpt_family="MIR"
repeat_region	116565..117029 /rpt_family="MLT1"
repeat_region	117948..118007 /rpt_family="MIR"
repeat_region	118345..118751 /rpt_family="THE1"

repeat_region /rpt_family="MIR"
119220..119507

Note: remainder of annotations omitted.

Query Match 73.9%; Score 17; DB 27; Length 177142;
Best Local Similarity 90.5%; Pred. No. 1.33e+01;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 77936 AGGCTGAACCCCTGTCCTGC 77956

Qy 3 aggtgaaacctgtcctgtc 23
|||||

RESULT 11
LOCUS HSU85195 200349 bp DNA PRI 27-MAY-1997
DEFINITION Homo sapiens BAC129, complete sequence.

ACCESSION U85195

NID 9121229

KEYWORDS HTG.

SOURCE human.

ORGANISM Homo sapiens

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Euthera; Primates; Catarrhini; Homiidae;
Homo.

REFERENCE 1 (bases 1 to 200349)

AUTHORS Boysen,C., Simon,M.I. and Hood,L.

TITLE Analysis of the 1.1-Mb human alpha/delta T-cell receptor locus with
bacterial artificial chromosome clones

JOURNAL Genome Res. 7 (4), 330-338 (1997)

MEDLINE 97264339

REFERENCE 2 (bases 1 to 200349)

AUTHORS Smith,T.M., Boysen,C., Simon,M.I. and Hood,L.

TITLE Direct Submission
Submitted (14-JAN-1997) Molecular Biotechnology, University of
Washington, Box 357730, Seattle, WA 98195, USA

JOURNAL Location/Qualifiers

FEATURES
source

1..200349

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 56864 a 41335 c 43211 g 58939 t

ORIGIN

Query Match 73.9%; Score 17; DB 28; Length 200349;

Best Local Similarity 100.0%; Pred. No. 1.33e+01;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 79275 TGAACCCCTGTCCTGC 79291

Qy 7 tgaacctgtcctgtc 23
|||||

RESULT 12
LOCUS HUA000658 250529 bp DNA PRI 02-SEP-1997
DEFINITION Homo sapiens T-cell receptor alpha delta locus from bases 1 to
250529 (section 1 of 5) of the Complete Nucleotide Sequence.

ACCESSION AE000658 AE000521

NID 92358019

KEYWORDS human.

SOURCE Homo sapiens

Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Euthera;
Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 250529)

AUTHORS Koop,B.F., Rowen,L., Wang,K., Kuo,C.L., Seto,D., Lenstra,J.A.,
Howard,S., Shan,W., Deshpande,P. and Hood,L.

TITLE The human T-cell receptor TCRAC/TCRDC (C alpha/C delta) region:
organization, sequence, and evolution of 97.6 kb of DNA

JOURNAL Genomics 19 (3), 478-493 (1994)

MEDLINE 94245236

REMARK This citation covers from bases 966383-1064019

REFERENCE 2 (bases 1 to 250529)

AUTHORS Boysen,C., Simon,M.I. and Hood,L.

TITLE Analysis of the 1.1-Mb human alpha/delta T-cell receptor locus with
bacterial artificial chromosome clones
JOURNAL Genome Res. 330, 330-338 (1997)
REFERENCE 3 (bases 1 to 250529)

AUTHORS Boysen,C., Inyol,L., Smith,T.M., Smit,A., Wang,K., Rowen,L. and
Hood,L.

TITLE T-Cell Receptor Alpha Delta Locus Complete Nucleotide Sequence

JOURNAL Unpublished (1997)

REMARK This citation covers bases 1-983545 and bases 1064020-1071650

REFERENCE 4 (bases 1 to 250529)

AUTHORS Boysen,C., Lee,I., Smith,T.M., Smit,A., Kai,W., Lee,R. and Leroy,H.

TITLE Direct Submission

JOURNAL Submitted (20-JUL-1997) Department of Molecular Biotechnology,
University of Washington, Box 357730, Seattle, Washington 98195,
USA

REMARK Complete nucleotide sequence of the human T-cell receptor alpha
delta locus

COMMENT Sequencing method: high redundancy shotgun. Interspersed Repeats
were identified with RepeatMasker (available from
http://ftp.genome.washington.edu/RM/RepeatMasker.html) Simple
sequence repeats were identified with sputnik (available from
http://serac.mbt.washington.edu/-chrisa/software/sputnik.html).

FEATURES
Location/Qualifiers

source

1..250529

/organism="Homo sapiens"

/db_xref="taxon:9606"

source

1..200349

/note="This span of sequence was submitted independently
as Genbank Accession Number U85195"

/organism="Homo sapiens"

/clone_lib="BAC129"

/clone="BAC129"

/db_xref="taxon:9606"

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

```

/rpt_type=tandem
/rpt_unit=AAAT
13164. .13433
/rpt_family="AluY"
13772. 14050
/rpt_family="AluJo"
14400. .14696
/rpt_family="AluSp"
14844. .15128
/rpt_family="AluJb"
15248. .15456
/rpt_family="AluJo"
15457. .15503
/rpt_family="AluJb"
complement(17422. .17722)
/rpt_family="AluSg"
18047. .18186
/rpt_family="L1ME3A"
complement(18243. .18564)
/rpt_family="AluSg"
complement(18565. .18838)
/rpt_family="AluSg"
complement(19843. .19905)
/rpt_family="AluSc"
20154. .20161
/rpt_type=tandem
/rpt_unit=AGG
20343. .20500
/rpt_family="MLT1G"
20899. .21196
/rpt_family="AluSg"
21206. .21323
/rpt_family="FLAM_A"
complement(21352. .21647)
/rpt_family="AluSg"
complement(22321. .22489)
/rpt_family="AluSx"
complement(22491. .22785)
/rpt_family="AluSg"
23022. .23132
/rpt_family="FLAM_A"
complement(23195. .23307)
/rpt_family="MER7B"
complement(23343. .23794)
/rpt_family="MER7B"
complement(23831. .24147)
/rpt_family="AluSx"
24989. .25266
/rpt_family="AluJb"
complement(25420. .25563)
/rpt_family="AluJo"
complement(25564. .25862)
/rpt_family="AluSg"
complement(25864. .25998)
/rpt_family="AluJo"
25999. .26056
/rpt_family="MIR"
complement(26111. .26311)
/rpt_family="AluJo"
33475. .33647
/rpt_family="FAM"
33648. .33940
/rpt_family="AluSg"
complement(34423. .34492)
/rpt_family="MIR"
35246. .35300
/rpt_family="MER5B"
35312. .35396
/rpt_family="FRAM"
35454. .35567
/rpt_family="AluJo"
35602. .35707
/rpt_family="MER5B"

```

```

repeat_region complement(36063. .36365)
/rpt_family="AluSp"
repeat_region complement(36431. .36689)
/rpt_family="AluJb"
repeat_region complement(37182. .37484)
/rpt_family="AluSp"
repeat_region complement(37489. .37806)
/rpt_family="AluSx"
37836. .37900
/rpt_family="L1MB8"
37836. .38013
/rpt_family="L1ME3"
37942. .38255
/rpt_family="L1MB4"
38258. .38432
/rpt_family="AluSg"
38460. .38760
/rpt_family="AluY"
38750. .38755
/rpt_type=tandem
/rpt_unit=AAAT
40842. .41142
/rpt_family="AluY"
41145. .41150
/rpt_type=tandem
/rpt_unit=AAAG
complement(41255. .41517)
/rpt_family="MIR"
complement(43371. .43461)
/rpt_family="MIR"
43583. .43716
/rpt_family="AluSc"
43717. .44018
/rpt_family="AluSx"

```

...
Note: remainder of annotations omitted.

Query Match 73.9%; Score 17; DB 27; Length 250529;
Best Local Similarity 100.0%; Pred. No. 1.33e+01;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 79275 TGAACCCCTGCTCTGC 79291

Qy 7 tgaacccctgcctgc 23

```

RESULT 13
LOCUS G05370 366 bp DNA STS 19-OCT-1993
DEFINITION human STS WI-7508.
ACCESSION G05370
NID 9852286
KEYWORDS STS sequence; primer; sequence tagged site.
SOURCE human Random genome wide STS created from sheared whole human DNA.
ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata;
Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primate;
Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 366)
AUTHORS Hudson,T.
TITLE Whitehead Institute/MIT Center for Genome Research; Random Genome
Wide STS
JOURNAL Unpublished (1995)
REFERENCE 2 (bases 1 to 366)
AUTHORS Hudson,T.
TITLE Whitehead Institute/MIT Center for Genome Research; Physically
Mapped STS
JOURNAL Unpublished (1995)
COMMENT
Contact: Thomas Hudson
Whitehead Institute/MIT Center for Genome Research
Whitehead Institute for Biomedical Research
9 Cambridge Center, Cambridge MA 02142 USA

```



```
misc_feature /note="encodes CRP domain"
2205. .2528
misc_feature /note="encodes scavenger receptor domain"
2529. .2621
misc_feature /note="encodes CRP domain"
2622. .3062
misc_feature /note="encodes Thr-rich domain"
3063. .3410
misc_feature /note="encodes CUB domain"
3411. .3758
misc_feature /note="encodes CUB domain"
3759. .3788
misc_feature /note="encodes Thr-Pro-rich domain"
3789. .4136
misc_feature /note="encodes CUB domain"
4137. .4484
misc_feature /note="encodes CUB domain"
4509. .4829
misc_feature /note="encodes scavenger receptor domain"
4830. .4889
misc_feature /note="encodes Ser-Thr-Pro-rich domain"
4890. .5225
misc_feature /note="encodes CUB domain"
5230. .6011
misc_feature /note="encodes 2P domain"
6135. .6203
misc_feature /note="encodes potential transmembrane domain"
6204. .6257
misc_feature /note="encodes potential cytoplasmic domain"
1580 a 1692 c 1748 g 1632 t
ORIGIN
Query Match 69.6%; Score 16; DB 29; Length 6652;
Best Local Similarity 90.0%; Pred. No. 5,26e+01;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 4441 CAAGGAGGCTTCAGGCT 4460
||||| ||||| ||||| ||
Cp 22 caaggacagggttcagcct 3

RESULT 15
LOCUS HPCJK04652 9461 bp ss-RNA VRL 09-FEB-1997
DEFINITION Hepatitis C virus isolate JK046 genomic RNA for polyprotein,
complete genome.
ACCESSION D63822
NID g1183030
KEYWORDS polyprotein.
SOURCE Hepatitis C virus (isolate:JK046) cDNA to genomic RNA.
ORGANISM Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepatitis C-like viruses.
REFERENCE 1 (sites)
AUTHORS Tokita,H., Okamoto,H., Iizuka,H., Kishimoto,J., Tsuda,F.,
Lesmana,I.A., Miyakawa,Y. and Mayumi,M.
Hepatitis C virus variants from Jakarta, Indonesia classifiable
into novel genotypes in the second (2e and 2f), tenth (10a) and
eleventh (11a) genetic groups
J. Gen. Virol. 77 (Pt 2), 293-301 (1996)
JOURNAL 96226020
MEDLINE 2 (bases 1 to 9461)
REFERENCE Okamoto,H.
AUTHORS Unpublished (1996)
REFERENCE 3 (bases 1 to 9461)
AUTHORS Okamoto,H.
TITLE Direct Submission
JOURNAL Submitted (10-Aug-1995) to the DDBJ/EMBL/GenBank databases. Hiroaki
Okamoto, Jichi Medical School, Immunology Division;
Minamikawachi-machi, Kawachi-gun, Tochigi 329-04, Japan
(Tel:0285-44-2111(ex.3334), Fax:0285-44-1557)
FEATURES
source Location/Qualifiers
1..9461
/organism="Hepatitis C virus"
```

```
misc_RNA
1..9461
/note="genomic RNA for polyprotein precursor, complete
genome"
/citation=[2]
/evidence=not_experimental
1..338
/citation=[2]
/evidence=not_experimental
339..9407
/citation=[2]
/codon_start=1
/evidence=not_experimental
/product="polyprotein"
/db_xref="PID:61010539"
/db_xref="PID:g1183031"
/translation="MSTNPKPQRTKRNTRRRPQDVKFFGGGOIVGGVYVLPGRGRL
GVRATKTSERSQPRGRQPIPKARRQTRGAWGQGYAMPGLYGNEGCGWAGLLSPRG
SRPTGPNDRRRSRNLKVIDITLCGLADLMGYIPVIGPLGGVAAALAHGVYAEAE
GVNYATGNLPGCSFSIFLLALLSLCTVPASAVNYANKSGIYHLTNDPNSMSVYAEAE
IILHLPGVCPIRTGNSQRCWTPATPTLAIPNSTVPASGFRQHLIDMLVGAALCSAMY
LGLDGGVFLVQLFTFRRIHQVQDCNSIITHVTHGRWAMDMWMSPTATFVW
SSALRAPQVDFIFAGGHWGIIIGALLYSTAANWAKVIIVLLLFAGVDASTVASSVS
QATSLVLSFASARQNLQINTNGSHINRTALNCNDSLOTGFTASLFRNKFNATG
CPELSACKTLDSDQGGPITYANISGPAVEKPYCWHYPPRPCEVVSALNRYGCVYC
FPPSPVLGTTDRRGNPTVTWCANETDFMSSLRCPAGGWGCTWMNTS3FVKTCGA
PCNIRPNPEENRTETLCPTDCFRKHGATYAKCGSGPLPRCLPYRLHLYPCS
TYNYLHKVRMYIAGSEHRTAACNWRGERCDLADRIEMSPLLFSTELAILPFC
FTMPALSTGLIHLQNVVDQYLYGLSTSYVNAIKWYVYVLRVAVITYGLNLW
LMLLIQAEALENIVLNAASAAATRGWECFLFMCWANYVVRVYVAVITYGLNLW
PLLLVLLPHRAYAYDVGAGSIIAALITFCVPHPSLGFVTKILLALIGPLYLLFASLLRPV
IEAKLHVWVPPFVRVGRDAIILLTCVPHPSLGFVTKILLALIGPLYLLFASLLRPV
VYVRAHALIRICALQNVNAGKYVQAAIILRAGSWGTGYIYDHLVPLTWASDGLRDLA
VAVPVPSPMEKKVITWGAADTAACGDIILAGLPVSARRGNLILLPADDPADKWSLL
APITAAQOGLLTITVSLTRDRKNEAGEIQLSTATQITFLATCVNGVYCWTVYHG
AGSKTLAPRGVPCVMTNVQDMYGPAGTRSYTCTCGASDLITLRQADVIPA
RRGDNRRAGLISPRITSLKSGGGLPCPSGHVVGFLFRAAVCTRGAKALDFVPCEA
MDATTSPFTDNSTPPAVPOAYQVYLHAPTSKGTSGKTKVPVAYASQGVKLVLPNSV
AATLSFGSVLSRAHGDIPNIRTVGRTITGAPITYSTYCKFLADGGCSGGAVDVILCD
ECHSTDPTVLGIGVLDQAEATAGRLVATATPPGSVTVPHNPQIETALPLTGEVP
FYGAALPEYIKGRRHLIFCSKKKCDLAAQRLTGLNVAFYRGVDVSVIPTSGDV
VVCATDALMTGYTDFSDVIDCNVATQVDFSLDPTFSETTTPQDANVRSQRGR
TGRGPGVRYVYVQSGRPSGMDFTVLCAYETGAAYELTTPAETVRLVRYLNTPL
PYCQDHLFEWEAVFTGLHIDAHFSLQKOGGENFAYLVAYQATCARAKAPPSMDT
MWKCLLRKPLTLPTGLLYRLGAVQNEVTPTPHVTKYIMACSADLEVTSTWVAG
GILAAIAAYCLTVGVSIVCGRITTSRPVAVIPDREVHQOYDEMEECSRHLPYVEGQ
QUAEQFKQVGLIQVTTQAEELKAPVHSAWPKLEQFQWKHMNFISGIIQLAGLST
IGSIGKVIYDILAGYAGVSGALVAFKIMSGETPSVEDMWNLLPALISGALVVG
VRAAILRRHVGPSEGAOWMNLIAFASGRNHVSPTHVVPETDASRAVNTLSLSTIT
SILRLKHLWITEDYATPCGSTRLDIMQVCTVLSDFVWLKASMLPSLPVPPFSCQ
RGYRTWGRDGCINTCCGASIAHVNKRWIRVGRPCSNVWNGTFPNTATTGPS
IPIPAIPYKALWRVSAIEYVEVVRSHYITGTAENTKCPVAPPEFTEEDGV
RUHRVAPCEKPLTRDEVTFTVGLTYVVGSQLPEPDLVTVSMRDPHPTAEAE
SRRLKRGSPSLASSQSLSAKATCTTHADHPDAELVEANLLRMKMGNTIRV
ESENKIVILDSPELKAEDFDRREISVAAEHRPPRFKYPALVPWARPDPNPLELFW
KAPYDPPVVGSCALPPQGLPVPPRRKKLVQDDSVVGHVLAOLAEKSFATPDOP
QTNDSGHTNGAASLPAEDDDADSDSSMPPLEGEPPGDDLSDDGSSGSSWSTVS
SEETSVCCSMSYSWTGALITPCAAEEBKLPISPLSNTLIRHNHNSYSSAALRQ
KKVTFDQVVDQHYDYLKEMKARSATSKLLSEACDLTPAHSAKSKFGYCAKD
VRGRTSKALNHINSYWDLEDNVTPIPTTINAKNEVFCVDVSKGRKPARLIVYDPL
SVRCEKRALYDVRKLPVAVMGAAVQYSPORVEYLLKIKWSKKTPMGFSYDTRC
FDSVTYERDITREESIYOCCELDPVARKAISLTERLYGGPMYSQSGCCYRBCRA
SGVLPSTMGNTLTCYKAKAKAAGLNFMVCGDDLVLVISELGSVEDSASALRAF
DAMTRYSGAPGDEPHEPVDLEHITSCSNVSAHDHTGQRIYIYITRDPDYNLARAAM
EPARTVNSYMGNTINMYAPTIVRWVLMTHFFGLQPOEQLHKLDFMTDVTGVTNTV
PLDLQIIRLHGMAFSLHGYSPGELRVGACLRKLAGPLRARHRRARAVRAKLLIA
QGGKAAICGMVLFNVAVKTKLTLPLRDAHRDLDSGWFWAGVYSGGDIHFHSVSHPRV
LILCLLLLTGVGVGIFLPPR"
```

```
9408..9461
/note="with 22nt of T-stretch"
/citation=[2]
```

3'UTR

BASE COUNT 2004 a 2729 c 2587 g 2141 t
ORIGIN /evidence-not_experimental

Query Match 69.6%; Score 16; DB 34; Length 9461;
Best Local Similarity 94.4%; Pred. No. 5.26e+01;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 4124 GGCTGCAACCCCTGTCCTT 4141
|||||
Qy 4 ggctgaaccctgtcctt 21

Search completed: Thu Apr 1 05:24:56 1999
Job time : 363 secs.

THIS PAGE BLANK (USPTO)

WIREH

(TM)

Release 3.1A John F. Collins, Blocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_n n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Thu Apr 1 05:30:47 1999; MasPar time 9.72 Seconds
Tabular output not generated. 321.923 Million cell updates/sec

Title: >US-09-238-972-2
Description: (1-23) from US09238972.seq
Perfect Score: 23 1 gtagcgtgaacacctgtctctgc 23
N.A. Sequence: catccgacttgggacaggaacg
Comp:

Scoring table: TABLE default
Gap 10

Nmatch STD : Dbase 0; Query 0

Searched: 188442 seqs, 68026449 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40

Statistics: Mean 5.361; Variance 3.030; scale 1.769

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
C 1	23	100.0	2397	11	Q63614 T-cell transmembrane	3.95e-04
C 2	22	95.7	2397	3	Q14533 Tea gene (CDNA 20.5).	1.61e-03
C 3	19	82.6	91	9	Q51746 Oligonucleotide probe	9.75e-02
C 4	18	78.3	91	9	Q51746 Oligonucleotide probe	3.68e-01
C 5	17	73.9	523	10	Q58556 Mercury resistant con	1.35e+00
C 6	16	69.6	204	1	N81164 Base substituted E.co	4.85e+00
C 7	15	65.2	523	10	Q58555 Mercury resistant con	1.69e+01
C 8	14	60.9	162	10	Q57461 Ribosome protein L6	1.68e+01
C 9	14	60.9	201	40	V24608 H. pylori ORF 01c116	5.68e+01
C 10	14	60.9	408	32	T59503 Humanised mouse anti-	5.68e+01
C 11	14	60.9	486	21	T27925 PRLTS gene exon 3 reg	5.68e+01
C 12	14	60.9	486	22	T27168 PRLTS coding sequence	5.68e+01
C 13	14	60.9	1330	14	Q75928 Mouse oploid receptor	5.68e+01

14	14	60.9	1452	34	T89585	Rat orphanin FQ recep	5.68e+01
15	14	60.9	1452	34	T90381	Rat methadone-specifi	5.68e+01
C 16	14	60.9	1502	21	T27922	PDGF receptor beta-11	5.68e+01
C 17	14	60.9	1502	23	T27165	PRLTS coding sequence	5.68e+01
C 18	14	60.9	1509	11	O64910	Thermus thermophilus	5.68e+01
C 19	14	60.9	1567	14	O89233	Rat oploid receptor c	5.68e+01
C 20	14	60.9	2224	15	O89544	P75 Tumour Necrosis F	5.68e+01
C 21	14	60.9	2339	2	Q10956	Encodes human 75kD TN	5.68e+01
C 22	14	60.9	2393	2	Q10907	40kD TNF inhibitor pr	5.68e+01
C 23	14	60.9	2443	33	T85094	Mouse transforming gr	5.68e+01
C 24	14	60.9	2600	15	Q90096	Mouse kappa-3 oploid	5.68e+01
C 25	14	60.9	2706	15	Q92972	Rat opiorph receptor	5.68e+01
C 26	14	60.9	3226	9	Q52675	F-spondin coding sequ	5.68e+01
C 27	14	60.9	3607	31	T43498	ATM gene 3' untransla	5.68e+01
C 28	14	60.9	3607	37	T45420	3' UTR of the ataxia-	5.68e+01
C 29	14	60.9	5100	15	Q90769	BAV3 E3 and fiber gen	5.68e+01
C 30	14	60.9	5912	31	T43519	ATM gene clone 7-9.	5.68e+01
C 31	14	60.9	5912	37	T45423	Partial ATM gene tran	5.68e+01
C 32	14	60.9	8806	36	T85392	Human TRIO phosphopro	5.68e+01
C 33	14	60.9	9515	14	Q55145	Pseudomonas aeruginos	5.68e+01
C 34	14	60.9	9520	2	N91779	Sequence encoding mou	5.68e+01
C 35	14	60.9	12778	31	T68784	ATM mutant G9170C cod	5.68e+01
C 36	14	60.9	17041	3	Q21065	Genomic DNA of human	5.68e+01
C 37	14	60.9	17041	12	Q67057	PACAPAP38 DNA.	5.68e+01
C 38	14	60.9	53526	40	T94101	Human PKD1 gene.	5.68e+01
C 39	14	60.9	53577	40	T94108	Human PKD1 locus betw	5.68e+01
C 40	14	60.9	53577	28	T18551	Human polycystic kidn	5.68e+01
C 41	13	56.5	105	33	T60577	K-coil peptide DNA.	1.84e+02
C 42	13	56.5	1745	40	V15235	Class I acyl-ACP thio	1.84e+02
C 43	13	56.5	1930	40	V16738	CDNA encoding human C	1.84e+02
C 44	13	56.5	2376	21	T28998	N. meningitidis serot	1.84e+02
C 45	13	56.5	3196	40	V20420	Human discs large1 g	1.84e+02

ALIGNMENTS

RESULT 1
ID Q63614 standard; cDNA to mRNA; 2397 BP.
AC Q63614;
DT 06-JAN-1995 (first entry)
DE T-cell transmembrane protein 20.5 cDNA.
KW transmembrane protein; T-lymphocytes; Tea gene; manipulation;
KW T-cell development; altering tumorigenic phenotype;
OS antibody production; identify and target drugs; ds.
Mus musculus.
FH Key Location/Qualifiers
FT misc_signal 407..413
FT /tag= a
FT /standard_name= Kozaks consensus sequence
FT cds 410..1771
FT /tag= b
FT /product= T cell transmembrane protein
FT /note= "Tea gene product (clone 20.5)"
PN US5312733-A.
PD 17-MAY-1994.
PF 13-APR-1990; 509684.
PR 13-APR-1990; US-509684.
PR 11-APR-1991; US-686322.
PA (RERE-) RES DEV FOUND.
PI Macleod CL;
DR WPI; 94-159119/19.
DR P-PSDB; R53467.
DR Expression vectors contg T-cell DNA sequences - coding for T-cell
PT trans membrane proteins
PS Claim 1; Fig 19A-19D: 56pp; English.
CC Q63614 encodes a novel T-cell multiple transmembrane protein (R53467).
CC The novel gene, 20.5, also referred to as Tea, is inducible in
CC SL 12.4 cells (CD4 CD8 double negative phenotype, therefore
CC resembling thymocytes at a relatively early stage of development.).
CC The Tea gene has been mapped to murine Chromosome 8. The cDNA does
CC not contain a polyadenylation signal sequence or a poly A tract. The
CC tea gene is differentially expressed in T lymphoma cells and activated
CC T lymphoid cells from normal spleen. The T cell protein may be useful

CC for manipulating regulation of T cell development, altering the
 CC tumorigenic phenotype and localising metastatic foci of tumours, and for
 CC producing antibodies which may be used to identify or target drugs
 CC or other agents to specific cell types expressing the proteins.
 SQ Sequence 2397 BP; 548 A; 571 C; 587 G; 691 T;

Query Match 100.0%; Score 23; DB 11; Length 2397;
 Best Local Similarity 100.0%; Pred. No. 3.95e-04;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1201 gcaaggacagggttcagcctac 1223
 Cp 23 gcaaggacagggttcagcctac 1

RESULT 2
 ID Q14533 standard; DNA; 2397 BP.
 AC Q14533;
 DT 29-JAN-1992 (first entry)
 DE Tea gene (cDNA 20.5).
 KW Multiple membrane spanning protein; T cell; development; ss.
 OS Mus musculus.
 FH Key Location/Qualifiers
 FT cds 410..1767
 FT /*tag= a
 PN WO9116430-A.
 PD 31-OCT-1991.
 PF 12-APR-1990; U02518.
 PR 13-APR-1990; US-509684.
 PA (REDE-) RES DEV FOUND.
 PI Macleod CL;
 DR WPI; 91-339818/46.
 DR P-PSDB; R14645.
 PT New recombinant polypeptide comprising a T-cell protein - used to
 PT regulate T-cell development and tumorigenic phenotype and to
 PT block T-cell activation in auto-immune disease
 PS Disclosure; Fig 13; 98pp; English.
 CC The 23 Ns in the sequence represent bases illegible in the
 CC specification.
 CC The 20.5 gene, also referred to as Tea, identifies transcripts
 CC found in only a limited number of tissues. Tea transcripts are
 CC induced in splenocytes activated with the T cell Mitogen ConA.
 CC Unlike other known genes expressed in activated T cells, the Tea gene
 CC appears to encode a protein which traverses the membrane multiple times,
 CC whereas the large number of known integral membrane protein which are
 CC induced in T cell activation are single membrane spanning proteins.
 CC See also Q14530-34.
 SQ Sequence 2397 BP; 541 A; 570 C; 579 G; 684 T;

Query Match 95.7%; Score 22; DB 3; Length 2397;
 Best Local Similarity 100.0%; Pred. No. 1.61e-03;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1202 caaggacagggttcagcctac 1223
 Cp 22 caaggacagggttcagcctac 1

RESULT 3
 ID Q51746 standard; cDNA; 91 BP.
 AC Q51746;
 DT 31-MAY-1994 (first entry)
 DE Oligonucleotide probe MK14-A
 KW Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
 OS ss.
 FH Key Location/Qualifiers
 FT -35_signal 61..66
 FT /*tag= a
 FT -10_signal 84..90
 FT /*tag= b
 FT rbs 101..108
 FT /*tag= c
 FT cds 116..523
 FT /*tag= d
 FT /product= merR(3) gene product.

PT New oligo:nucleotide probes specific for Mycobacteria - used for
 PT detection and amplification of Mycobacteria nucleic acid in
 PT samples
 PS Claim 3; Page 14; 23pp; English.
 CC Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
 CC (Q51735). It hybridized to all spp. of mycobacteria tested, but
 CC cross reacted to a few non-mycobacterial spp. The probe may
 CC be useful as an initial screen for mycobacterial infection.
 CC See also Q51735-45 and Q51747-59.
 SQ Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;

Query Match 82.6%; Score 19; DB 9; Length 91;
 Best Local Similarity 0.0%; Pred. No. 9.75e-02;
 Matches 0; Conservative 21; Mismatches 2; Indels 0; Gaps 0;

Db 31 vvhvhvsvvvvhvvhvvhvvhv 53
 Cp 23 gcaaggacagggttcagcctac 1

RESULT 4
 ID Q51746 standard; cDNA; 91 BP.
 AC Q51746;
 DT 31-MAY-1994 (first entry)
 DE Oligonucleotide probe MK14-A
 KW Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
 OS ss.
 OS Synthetic.
 PN EP-571911-A.
 PD 01-DEC-1993.
 PF 24-MAY-1993; 108325.
 PR 26-MAY-1992; US-889651.
 PA (BECT) BECTON DICKINSON CO.
 PI Shank DD, Spears PA;
 DR WPI; 93-378844/48.
 PT New oligo:nucleotide probes specific for Mycobacteria - used for
 PT detection and amplification of Mycobacteria nucleic acid in
 PT samples
 PS Claim 3; Page 14; 23pp; English.
 CC Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
 CC (Q51735). It hybridized to all spp. of mycobacteria tested, but
 CC cross reacted to a few non-mycobacterial spp. The probe may
 CC be useful as an initial screen for mycobacterial infection.
 CC See also Q51735-45 and Q51747-59.
 SQ Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;

Query Match 78.3%; Score 18; DB 9; Length 91;
 Best Local Similarity 0.0%; Pred. No. 3.68e-01;
 Matches 0; Conservative 20; Mismatches 2; Indels 0; Gaps 0;

Db 25 hshvhhvvhvvhvsvvvvhvvhv 46
 Qy 2 taggctgaacctgtcctgtgc 23

RESULT 5
 ID Q58556 standard; DNA; 523 BP.
 AC Q58556;
 DT 16-SEP-1994 (first entry)
 DE Mercury resistant control gene merR(3).
 KW Resistance; mercury; selectable marker; Thiobacillus ferrooxidans;
 KW transformation; detection; ss.
 OS Thiobacillus ferrooxidans.
 FH Key Location/Qualifiers
 FT -35_signal 61..66
 FT /*tag= a
 FT -10_signal 84..90
 FT /*tag= b
 FT rbs 101..108
 FT /*tag= c
 FT cds 116..523
 FT /*tag= d
 FT /product= merR(3) gene product.

Q58555 standard; DNA; 523 BP.

Q58555; AC 16-SEP-1994 (first entry)

DE Mercury resistant control gene merR(2).

KW Resistance: mercury; selectable marker; Thiobacillus ferrooxidans;

KW Transformation; detection; ss.

TH Thiobacillus ferrooxidans.

OS Thiobacillus ferrooxidans.

Key Location/Qualifiers

FT -35_signal 61..66 /*tag= a

FT -10_signal 84..89 /*tag= b

FT rbs 101..107 /*tag= c

FT cds 116..523 /*tag= d

FT /product= merR(2) gene product.

PN J06000083-A.

PD 11-JAN-1994.

PE 17-JAN-1991; 018338.

PR 17-JAN-1991; JP-018338.

PA (AKIT-) AKITA KEN.

PA (DOWA) DOWA MINING CO LTD.

DR WPI; 94-077131/10.

DR P-PSDB; R49669.

PT Mercury resistant control gene merR and shuttle vector - for enhanced expression of mercury resistance marker in transformed Thiobacillus sp.

PT Claim 1; Page 3; 26pp; Japanese.

PS The mercury resistance genes can be used as selectable markers when used to transform other bacteria.

CC Sequence 523 BP; 118 A; 129 C; 152 G; 124 T;

SO Query Match 65.2%; Score 15; DB 10; Length 523; Best Local Similarity 89.9%; Pred. No. 1.69e-01; Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 285 aaggactgggcttcagcct 303
||||| ||| |||||||

Cp 21 aaggacagggttcagcct 3

RESULT 8

ID Q57461 standard; cDNA to mRNA; 162 BP.

AC Q57461;

DT 19-OCT-1994 (first entry)

DE Ribosome protein L6 like protein.

KW Human cDNA; library; enzyme; protein; ss.

KW Homo sapiens.

PN W09403599-A.

PD 17-FEB-1994.

PF 04-AUG-1993; J01095.

PR 04-AUG-1992; JP-208077.

PR 13-NOV-1992; JP-327619.

PR 26-FEB-1993; JP-061431.

PA (SAGA) SAGAMI CHEM RES CENTRE.

PI Iwahori A, Kato S, Kato T, Kim N, Oh S, Sekine S;

DR WPI; 94-065688/08.

DR P-PSDB; R46122.

PT cDNA of human origin and proteins coded by it - which may be expressed by in vivo or in vitro translation using sense RNA or antisense DNA corresponding to the cDNA.

PT Claim 1; Page 83; 167pp; Japanese.

CC mRNA expressed in human fibrosarcoma cell line HT-1080 was isolated and used to construct a cDNA library using vector pKAL1. Clone HP00212 encoding Ribosome protein L6-like protein was isolated.

CC Sequence 162 BP; 59 A; 32 C; 47 G; 24 T;

SO Query Match 60.9%; Score 14; DB 10; Length 162; Best Local Similarity 88.9%; Pred. No. 5.68e-01; Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Db 134 gccgaaccctgtctctt 151
    ||| ||||| |||||
Oy 5 gctgaaccctgtctctt 22

RESULT 9
ID V24608 standard; DNA; 201 BP.
AC V24608;
DT 15-JUN-1998 (first entry)
DE H. pylori ORF 01cel1618orf3.
KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW identification; binding compound; bacteria; life cycle; activator;
KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis; ds.
OS Helicobacter pylori.
FH Key Location/Qualifiers
FT 1..201
FT CDS
FT /tag= a
FT /note= "no stop codon given"
FN W09737044-A1.
PD 09-OCT-1997.
PF 27-MAR-1997; U05223.
PR 06-DEC-1996; US-761318.
PR 29-MAR-1996; US-625811.
PR 02-APR-1996; US-758731.
PR 25-OCT-1996; US-736905.
PR 28-OCT-1996; US-738859.
PA (ASTR ) ASTRA AB.
PI Alm RA, Smith D.
DR WPI: 97-503122/46.
DR P-PSDB: W53199.
DT Helicobacter pylori nucleic acid sequences and encoded
PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori
PT infection and for diagnosis of H. pylori infection
PS Claims 5.6; Page 137; 1145pp; English.
CC This sequence encodes a H. pylori protein of unspecified function.
CC Infection or to identify H. pylori polypeptide binding compounds,
CC useful as potential H. pylori life cycle activators or inhibitors. The
CC DNA and probes derived from it may be used for the identification of
CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic
CC acid sequences complementary to the DNA act as antisense sequences and
CC can be used to prevent the translation of H. pylori mRNA. Antibodies
CC against the protein can be used in immunoassays to evaluate the abundance
CC and distribution of H. pylori-specific antigens. The genomic sequence of
CC H. pylori (ATCC 55679) was determined from overlapping contigs generated
CC by mechanically shearing the bacterial DNA. The sequences were analysed
CC for ORF of at least 180 nucleotides, and the predicted coding regions
CC defined by computer evaluation. To identify likely H. pylori antigens for
CC vaccine development, the amino acid sequences predicted from various ORF
CC were analysed for significant homology to other known or exported
CC membrane proteins. Having identified and determined the sequences of
CC interest, particular regions can be isolated from H. pylori by PCR
CC amplification for recombinant polypeptide production, e.g. in E. coli
CC hosts.
SQ Sequence 201 BP; 73 A; 38 C; 32 G; 58 T;

Query Match 60.9%; Score 14; DB 40; Length 201;
Best Local Similarity 93.8%; Pred. No. 5.68e+01;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 89 agcctgaacccttctt 104
    ||||| ||||| ||
Oy 3 agcctgaaccctgtc 18

RESULT 10
ID T59503 standard; cDNA; 408 BP.
AC T59503;
DT 20-OCT-1997 (first entry)
DE Humanised mouse anti-human Fas ligand antibody F919 heavy chain cDNA.
KW Human; Fas ligand; antigen; neutral; antibody; apoptosis; HIV;
KW induction; assay; enzyme linked immunosorbant assay; diagnosis;
KW disease; hepatitis B; hepatitis C; human immunodeficiency virus;

graft versus host disease; ulcerative colitis; sequelae; chimeric;
myocardial infarction; mouse; murine; monoclonal; treatment; ss.
Chimeric - Mus spp.
Chimeric - Homo sapiens.
Synthetic.
FH Key Location/Qualifiers
FT mat_peptide 1..408
FT /tag= a
FN W09702290-A1.
PD 23-JAN-1997.
PF 01-JUL-1996; J01820.
PR 17-MAY-1996; US-649100.
PR 30-JUN-1995; JP-188480.
PA (MOCH ) MOCHIDA PHARM CO LTD.
PA (OSAB-) OSABA BIOSCIENCE INST.
PI Co MS, Matsusue T, Nagata S, Shirakawa K, Vasquez M;
DR WPI: 97-108917/10.
DR P-PSDB: W11818.
DT Antibody reactive with Fas ligand capable of inducing apoptosis -
PT used for diagnostic assay of Fas ligand in body fluids and for
PT treatment of diseases in which Fas ligand/Fas antigen is involved
PS Disclosure; Fig 13; 164pp; Japanese.
CC The present sequence encodes the heavy chain of the humanised
CC mouse anti-apoptosis inducing human Fas ligand, monoclonal
CC antibody, F919. The antibody can be used in a Fas ligand assay,
CC e.g. an enzyme linked immunosorbant assay, to diagnose diseases
CC in which the Fas ligand/Fas antigen system is implicated e.g.
CC hepatitis B/C, human immunodeficiency virus, graft/host disorders,
CC ulcerative colitis or sequelae of myocardial infarction. The
CC antibody may also be used to treat such diseases.
SQ Sequence 408 BP; 100 A; 95 C; 110 G; 103 T;

Query Match 60.9%; Score 14; DB 32; Length 408;
Best Local Similarity 100.0%; Pred. No. 5.68e+01;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 180 agcacagggtttca 193
    ||||| ||||| ||
Cp 20 agcacagggtttca 7

RESULT 11
ID T27925 standard; DNA; 486 BP.
AC T27925;
DT 24-AUG-1996 (first entry)
DE PRLTS gene exon 3 region.
KW Platelet-derived growth factor receptor beta-like tumour suppressor;
KW PDGF; PRLTS; carcinogenesis; tumorigenesis; lung cancer;
KW hepatocellular carcinoma; colorectal cancer; diagnosis; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT misc_difference 46
FT /tag= a
FT /note= "base n at position 46 is not identified
FT in the specification"
FT misc_difference 48
FT /tag= b
FT /note= "base n at position 48 is not identified
FT in the specification"
FT primer_bind complement (49..71)
FT /tag= c
FT /note= "primer for exon 3"
FT exon 74..371
FT /tag= d
FT /label= Exon-3
FT misc_difference 376
FT /tag= e
FT /note= "base n at position 376 is not identified
FT in the specification"
FT misc_difference 391
FT /tag= f
FT /note= "base n at position 391 is not identified
FT in the specification"

```



```

FT primer_bind 393..415
FT /*tag= g
FT /note= "primer for exon 3"
PN EP-714981-A2.
PD 05-JUN-1996
PF 26-JUL-1995; 111769.
PR 29-JUL-1994; JP-178131.
PA (CANC-) CANCER INST.
PA (EISA ) EISAI CO LTD.
PI Fujiwara Y, Nakamura Y;
DR WPI: 96-269714/28.
DR P-PSDB: R95934.
DR PDGF-receptor beta-like tumour suppressor protein - for detecting
FT gene mutation(s) to diagnose, monitor, etc. cancers of lung, liver
FT or colon
PS Claim 11; Page 36; 49pp; English.
CC 7 Genomic DNA fragments (T27923-29) were detected in cosmids covering
CC a deletion region in the vicinity of p21.3 to 22 of human chromosome 8.
CC This region is commonly deleted in cases of lung cancer, hepatocellular
CC carcinoma and colorectal cancer. The DNA fragments are derived from a
CC novel platelet-derived growth factor receptor beta-like tumour
CC suppressor protein (PRLTS) gene; a cDNA sequence is given in T27922.
CC Exon 3 (T27925) codes for an internal portion (R95934) of PRLTS
CC protein (see also R95932). Primers (see also T27932-33) designed to
CC amplify exon 3 were used for single strand polymorphism analysis of
CC cancer samples.
CC (N.B. the genomic DNA fragments are called 'introns' in the Claims,
CC but 'exons' elsewhere in the specification).
CC Sequence 486 BP; 149 A; 109 C; 96 G; 128 T;
SQ
Query Match 60.9%; Score 14; DB 21; Length 486;
Best Local Similarity 88.9%; Pred. No. 5.68e+01;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 175 aagcacagggttcagcc 192
Cp 21 aagcacagggttcagcc 4

RESULT 12
ID T27168 standard; DNA; 486 BP.
AC T27168;
DT 19-NOV-1996 (first entry)
DE PRLTS coding sequence, exon 3.
KW Human; platelet-derived growth factor; chromosome 8; deletion; PRLTS;
KW liver cancer; liver non-small cell cancer; hepatocellular carcinoma;
KW PDGF receptor beta-like tumour suppressor protein; colon cancer; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT intron 1..73
FT /*tag= a
FT exon 74..371
FT /*tag= b
FT /number= 3
FT /note= "Encodes PRLTS amino acids 19-118, see also
FT R96129"
FT intron 372..486
FT /*tag= c
PN J08092291-A.
PD 09-APR-1996.
PF 06-JUN-1995; 139111.
PR 29-JUL-1994; JP-178131.
PA (EISA ) EISAI CO LTD.
PA (GANK-) ZH GAN KENYUKAI.
DR WPI: 96-236101/24.
DR PRLTS protein and a DNA encoding it - used in the detection and
FT treatment of cancer
PS Claim 8; Page 12; 18pp; Japanese.
CC The sequences given in T27166-72 are genomic DNA fragments which
CC contain the human platelet-derived growth factor receptor beta-like
CC tumour suppressor protein (PRLTS) gene exons. PRLTS is the product of
CC a gene on chromosome 8 present in a region commonly deleted in cases
CC of liver cancer, liver non-small cell cancer and colon cancer. The

```

```

CC PRLTS protein can be used as a research agent, a detecting and
CC diagnosing reagent and in the treatment of cancer.
SQ Sequence 486 BP; 149 A; 109 C; 96 G; 128 T;

Query Match 60.9%; Score 14; DB 22; Length 486;
Best Local Similarity 88.9%; Pred. No. 5.68e+01;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 175 aagcacagggttcagcc 192
Cp 21 aagcacagggttcagcc 4

RESULT 13
ID Q75928 standard; DNA; 1330 BP.
AC Q75928;
DT 17-AUG-1995 (first entry)
DE Mouse opioid receptor-like receptor MOP2 cDNA.
KW Mouse; kappa; delta; mu; opioid receptor; brain; primer; PCR; amplify;
KW transmembrane domain; somatostatin; receptor; human; expression vector;
KW truncate; chimaeric; assay; probe; ss.
OS Mus musculus.
FH Key Location/Qualifiers
FT cds 161..1264
FT /*tag= a
FT /product= mouse opioid receptor-like receptor
PN W09428132-A.
PD 08-DEC-1994.
PF 20-MAY-1994; U05747.
PR 20-MAY-1993; US-066296.
PR 30-JUL-1993; US-100694.
PR 05-NOV-1993; US-147592.
PA (ARCH-) ARCH DEV CORP.
PI Bell GI, Reisine T, Yasuda K;
DR WPI: 95-022804/03.
DR P-PSDB: R67671.
DR Polynucleotides and peptides derived from opioid receptor
FT polypeptides - for use in therapeutic compositions and in
FT screening assays for useful drug substances.
PS Claim 28; Page 225-229; 300pp; English.
CC The nucleotide sequence of the novel mouse opioid receptor-like receptor
CC gene MOP2. MOP2 is a mouse receptor with pharmacological properties which
CC are dissimilar to the properties of classic opioid receptors such kappa,
CC delta, mu or sigma. It has been found that drug of high abuse potential
CC or analgesic potency bind selectively to this receptor. This suggests
CC that this receptor could be important in the development of drugs to
CC treat addiction. Other opioid receptors isolated and produced such as the
CC novel mouse kappa and delta opioid receptors (Q75926-7) are useful for
CC the development of novel assays designed to select or improve substances,
CC capable of interacting with the opioid receptor proteins, for use in
CC diagnosis, drug design and therapeutic applications.
CC Sequence 1330 BP; 250 A; 368 C; 350 G; 362 T;
SQ

Query Match 60.9%; Score 14; DB 14; Length 1330;
Best Local Similarity 85.0%; Pred. No. 5.68e+01;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 436 ggctgataccctggtctgc 455
Cp 4 ggctgaaccctgtcctgc 23

RESULT 14
ID T89585 standard; cDNA; 1452 BP.
AC T89585;
DT 12-JAN-1998 (first entry)
DE Rat orphanin FQ receptor cDNA clone LCI32.
KW Orphanin FQ receptor; binding; locomotor disease; diagnosis;
KW treatment; opioid inhibitor; opiate induced hypothermia; drug design;
KW morphine induced analgesia; methadone specific opioid receptor; ss.
OS Rattus sp.
FH Key Location/Qualifiers
FT 5'UTR 1..181

```

```

FT CDS /*tag= a
FT 182..1285
FT /*tag= b
FT 1286..1452
FT /*tag= c
PN WO9707212-A1.
PD 27-FEB-1997.
PR 12-AUG-1996; UI3305.
PR 03-NOV-1995; US-553058.
PR 11-AUG-1995; US-514451.
PA (UYOR-) UNIV OREGON HEALTH SCI.
PI Bunzow JR, Civelli O, Grandy DK, Grisel JE, Mogil JS;
PI Monnsma FJ, Nothacker H-P, Reinscheid RK;
DR WPI; 97-165296/15.
DR P-PSDB; W25217.
PT Mammalian orphanin FQ receptor inhibitor - used in the diagnosis and
PT treatment of locomotor disease.
PS Claim 5; Page 42-43; 68pp; English.
CC T89585 is rat orphanin FQ (OFQ) receptor cDNA clone LCI32. Highly
CC specific peptides that bind the OFQ receptor were identified as
CC OFQ receptor inhibitors. The peptides can be used to antagonise a
CC physiological effect of an opioid in an animal. The peptides antagonise
CC opiate induced hypothermia and morphine induced analgesia in animals.
CC They may also be used in the diagnosis and treatment of locomotor
CC disease. The peptides may also be used in the design of a methadone
CC specific opioid receptor (MSOR), in drug design and for the isolation
CC of endogenous receptors for anti-opioid agonists and antagonists found
CC in plasma, serum, lymph, cerebrospinal fluid etc.
SQ Sequence 1452 BP; 276 A; 414 C; 377 G; 385 T;

```

Query Match 60.9%; Score 14; DB 34; Length 1452;
 Best Local Similarity 85.0%; Pred. No. 5.68e+01;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

Db 457 ggctgataccctggctcttcg 476
   ||||| ||||| |||||
OY 4 ggctgaacccctgtccttcg 23

```

```

RESULT 15
ID T90381 standard; cDNA; 1452 BP.
AC T90381;
DT 23-JAN-1998 (first entry)
DE Rat methadone-specific opioid receptor MSOR cDNA.
KW Methadone-specific opioid receptor; MSOR; rat; psychoactive drug;
OS Rattus sp.
FH Key Location/Qualifiers
FT CDS 182..1285
FT /*tag= a
PN US5658783-A.
PD 19-AUG-1997.
PF 08-NOV-1993; 149093.
PR 08-NOV-1993; US-149093.
PA (UYOR-) UNIV OREGON HEALTH SCI.
PI Bunzow JR, Civelli O, Grandy DK;
PI WPI; 97-424240/39.
DR P-PSDB; W26582.
PT New isolated mammalian methadone-specific opioid receptor gene -
PT used to develop products for screening for compounds which can be
PT used as psychoactive drugs
PS Claim 1; Column 17-20; 26pp; English.
CC This nucleotide sequence encodes a novel, claimed rat methadone-
CC specific opioid receptor (MSOR) (see W26582). It was isolated
CC from a rat brain library using degenerate primers (see T90382-83)
CC based on the putative 3rd and 7th transmembrane regions of a mouse
CC delta-opioid receptor. Also claimed are: (1) a hybridisation
CC probe; (2) a recombinant expression construct that is capable of
CC expressing the MSOR in eukaryotic or prokaryotic cell transformants;
CC and (3) a transformed cell culture that expresses the MSOR. The
CC probes can be used to determine the pattern, amount and extent of
CC expression of the MSOR gene in various tissues of mammals,
CC including humans, as well as in the detection and diagnosis of

```

```

CC genetic disease, or the detection of novel related receptor genes.
SQ Sequence 1452 BP; 276 A; 414 C; 377 G; 385 T;
Query Match 60.9%; Score 14; DB 34; Length 1452;
Best Local Similarity 85.0%; Pred. No. 5.68e+01;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Db 457 ggctgataccctggctcttcg 476
   ||||| ||||| |||||
OY 4 ggctgaacccctgtccttcg 23

```

Search completed: Thu Apr 1 05:32:50 1999
 Job time : 123 secs.

 WIP

 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPSrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
 Run on: Thu Apr 1 05:25:15 1999; Maspar time 88.29 Seconds
 Tabular output not generated. 466.521 Million cell updates/sec

Title: >US-09-238-972-2
 Description: (1-23) from US09238972.seq
 Perfect Score: 23
 N.A. Sequence: 1 gtagcgtgaaccccttcctgc 23
 Comp: catccgacttggacaggaacg

Scoring table: TABLE default
 Gap 10
 Nmatch STD : Dbase 0; Query 0
 Searched: 2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 0%
 Listing first 45 summaries
 Database: emb1-est56
 Database: 1:emb1 2:em_gss1 3:em_gss2 4:em_gss3
 5:gb_est1 6:gb_est10 7:gb_est11 8:gb_est12 9:gb_est13
 10:gb_est14 11:gb_est15 12:gb_est16 13:gb_est17
 14:gb_est18 15:gb_est19 16:gb_est2 17:gb_est20
 18:gb_est21 19:gb_est3 20:gb_est4 21:gb_est5 22:gb_est6
 23:gb_est7 24:gb_est8 25:gb_est9 26:gb_gss1 27:gb_gss2
 28:gb_gss3 29:gb_gss4

Statistics: Mean 6.917; Variance 1.457; scale 4.748
 Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
1	17	73.9	455 13	NR89c10.s1 NCI_CGAP_Pr
2	17	73.9	500 19	Yz39b02.s1 Homo sapien
3	17	73.9	554 13	zp31d05.s1 Stratagene
4	17	73.9	614 28	nbx50012101f CUGI Rice
5	16	69.6	194 20	Mouse expressed sequen
6	16	69.6	196 11	nz91g02.s1 NCI_CGAP_GC
7	16	69.6	248 14	ua62q11.s1 Soares 2NbM
8	16	69.6	251 28	HS_3050.B1.H03 MR CIT
9	16	69.6	270 20	F5516 Fetal heart Homo
10	16	69.6	274 21	mj34g06.r1 Soares mous
11	16	69.6	304 14	vc58b11.s1 NCI_CGAP_GC
12	16	69.6	320 11	vm53b05.s1 Stratagene
13	16	69.6	329 21	mb54h09.r1 Soares mous

c	14	16	69.6	342 23	AA065583	ml73g05.r1 Stratagene	3.94e-02
c	15	16	69.6	361 20	W00972	za55d07.r1 Soares feta	3.94e-02
c	16	16	69.6	367 7	AA511049	YH55h06.r1 Soares mous	3.94e-02
c	17	16	69.6	369 27	B86590	RPC111-24D20.TP RPC111	3.94e-02
c	18	16	69.6	384 11	AA084826	zn02d12.r1 Stratagene	3.94e-02
c	19	16	69.6	387 27	AQ054393	CIT-HSP-234015.1R CIT-	3.94e-02
c	20	16	69.6	395 21	W35803	mc14c04.r1 Soares mous	3.94e-02
c	21	16	69.6	399 23	AA105504	mm22a09.r1 Stratagene	3.94e-02
c	22	16	69.6	400 21	AA050861	me03f01.r1 Soares mous	3.94e-02
c	23	16	69.6	414 20	W64404	me03f01.r1 Soares mous	3.94e-02
c	24	16	69.6	417 11	AA726284	vu89e06.r1 Stratagene	3.94e-02
c	25	16	69.6	421 7	AA472878	vb30b10.r1 Soares mous	3.94e-02
c	26	16	69.6	421 18	AI155538	ue05c10.r1 Soares mous	3.94e-02
c	27	16	69.6	423 21	AA050482	mj20b01.r1 Soares mous	3.94e-02
c	28	16	69.6	458 13	AA806368	oc22e04.s1 NCI_CGAP_GC	3.94e-02
c	29	16	69.6	459 5	R20257	Yg19g11.r1 Homo sapien	3.94e-02
c	30	16	69.6	459 11	AA070324	zm58d08.r1 Stratagene	3.94e-02
c	31	16	69.6	466 23	AA218095	mv37c10.r1 Soares mous	3.94e-02
c	32	16	69.6	468 15	AI020661	uv97f02.r1 Soares mous	3.94e-02
c	33	16	69.6	469 21	AA041558	mj01g10.r1 Soares mous	3.94e-02
c	34	16	69.6	471 24	AA277613	va08b01.r1 Soares mous	3.94e-02
c	35	16	69.6	472 24	AA272778	va41h05.r1 Soares mous	3.94e-02
c	36	16	69.6	485 21	AA073883	mk02c09.r1 Soares mous	3.94e-02
c	37	16	69.6	491 21	AA064182	mj46f01.r1 Soares mous	3.94e-02
c	38	16	69.6	503 18	AI154798	uv59h12.r1 Soares mous	3.94e-02
c	39	16	69.6	508 23	AA250313	mw06c04.r1 Soares mous	3.94e-02
c	40	16	69.6	509 14	AA864963	oh36d06.s1 NCI_CGAP_Ki	3.94e-02
c	41	16	69.6	522 23	AA240820	mw24b05.r1 Soares mous	3.94e-02
c	42	16	69.6	524 24	AA267123	mz74e04.r1 Soares mous	3.94e-02
c	43	16	69.6	524 23	AA080271	mm34h09.r1 Stratagene	3.94e-02
c	44	16	69.6	559 27	AQ013928	RPC111-24C20.TKBR RPC1	3.94e-02
c	45	16	69.6	562 7	AA517515	vh92c09.r1 Knowles Sol	3.94e-02

ALIGNMENTS

RESULT 1
 LOCUS AA826760 455 bp mRNA EST 20-FEB-1998
 DEFINITION nr89c10.s1 NCI_CGAP_Pr25 Homo sapiens cDNA clone IMAGE:1175154
 similar to contains element MSRI repetitive element ;, mRNA
 sequence.
 AA826760
 EST
 G2900757
 KEYWORDS
 human.
 SOURCE
 Homo sapiens
 ORGANISM
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE
 1 (bases 1 to 455)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL
 Unpublished (1997)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Suzanne L. Topalian, M.D., Robert K. Bright,
 Ph.D.
 CDNA Library Preparation: Stratagene, Inc.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 429.
 Location/Qualifiers
 1. .455
 /organism="Homo sapiens"
 /note="Organ: prostate; Vector: Bluescript SK-; Site:1;
 ECORI; Site_2: XhoI; Cloned unidirectionally. Primer:

Oligo dT. Normal prostate epithelial cell line (HPV immortalized). 5' adaptor sequence: 5' GAATTCGGCAGAG 3', adaptor sequence: 5' CTCGAGTGTTCCTTTT 3', Average insert size: 1.1 kb.
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:1175154"
 /clone_lib="NCI_CGAP_Pr25"
 /tissue_type="epithelium (cell line)"
 /lab_host="SOLR (kanamycin resistant)"
 BASE COUNT 105 a 94 c 146 g 110 t
 ORIGIN

Query Match 73.9%; Score 17; DB 13; Length 455;
 Best Local Similarity 94.7%; Pred. No. 2.26e-03;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 424 GGACAGGGTTCAGCCTAC 442
 Cp 19 ggacagggttcagcctac 1
 ||||||| |||||||

RESULT 2
 LOCUS N66365 500 bp mRNA EST 08-MAR-1996
 DEFINITION Y23902.s1 Homo sapiens cDNA clone 285387 3', similar to gb:X63563
 DNA-DIRECTED RNA POLYMERASE II 140 KD POLYPEPTIDE (HUMAN);
 ACCESSION N66365
 NID g1218490
 KEYWORDS EST.
 SOURCE human clone-285387 primer-m13 -40 forward library=Worton Fetal
 Cochlea vector-pBluescript SK- host-SOLR cells (kanamycin resistant) Kslet-EcoRI Kslet2-XhoI The cDNA was oligo (dT) primed with an XhoI restriction enzyme recognition site and an 18 base poly dT sequence. For the 5' end, the synthesized cDNA termini were treated with T4 DNA polymerase and EcoRI adaptors were ligated to the blunt ends. adaptor linker: GAATTCGGCAGAG.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 500)
 AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevasik,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
 TITLE WashU-Merck EST Project
 JOURNAL Unpublished (1995)
 COMMENT Contact: Wilson RK
 WashU-Merck EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 High quality sequence stops: 273
 Source: IMAGE Consortium, LLNL
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Location/Qualifiers
 1..500
 /organism="Homo sapiens"
 /clone="285387"
 <1..>500

BASE COUNT 128 a 93 c 106 g 173 t
 ORIGIN

Query Match 73.9%; Score 17; DB 19; Length 500;
 Best Local Similarity 94.7%; Pred. No. 2.26e-03;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 319 GGTGAAGCCTGCTTCTG 337
 ||||||| |||||||

QY 4 ggcgtgaaccctgtctctg 22

RESULT 3
 LOCUS AA173270 554 bp mRNA EST 09-MAR-1998
 DEFINITION zp31d05.s1 Stratagene neuroepithelium (#937231) Homo sapiens cDNA
 clone 611049 3', mRNA sequence.
 ACCESSION AA173270
 NID g1753402
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 554)
 AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,W., Tan,F., Theising,B., White,Y., Wyllie,T., Waterston,R. and Wilson,R.
 TITLE WashU-NCI human EST Project
 JOURNAL Unpublished (1997)
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 1064 Std Error: 0.00
 Seq primer: -40M13 fwd. from Amersham
 High quality sequence stop: 433.
 Location/Qualifiers
 1..554
 /organism="Homo sapiens"
 /note="Vector: pBluescript SK-; Site1: EcoRI; Site2: XhoI; Cloned unidirectionally. Primer: Oligo dT. HT2 cells (Ntera-2/cl.D1) induced with Retinoic Acid for 24 hours. Average insert size: 1.5 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5' CTCGAGTGTTCCTTTT 3'."
 /db_xref="GDB:4626308"
 /db_xref="taxon:9606"
 /clone="611049"
 /clone_lib="Stratagene neuroepithelium (#937231)"
 /dev_stage="Ntera-2/RA neuroepithelial cells"
 /lab_host="SOLR (kanamycin resistant)"
 BASE COUNT 135 a 105 c 157 g 155 t 2 others
 ORIGIN

Query Match 73.9%; Score 17; DB 13; Length 554;
 Best Local Similarity 94.7%; Pred. No. 2.26e-03;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 432 GGACAGGGTTCAGCCTAC 450
 Cp 19 ggacagggttcagcctac 1
 ||||||| |||||||

RESULT 4
 LOCUS AQ158758 614 bp DNA GSS 10-SEP-1968
 DEFINITION nbxb0012101f CUGI Rice BAC Library Oryza sativa genomic clone
 nbxb0012101f, genomic survey sequence.
 ACCESSION AQ158758
 NID g3555783
 KEYWORDS GSS.
 SOURCE Oryza sativa.
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Charophyta/Embryophyta group; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
 REFERENCE 1 (bases 1 to 614)

AUTHORS Yu, Y., Budiman, M.A., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R., Phillips, K., Sasinowski, M., Dean, R.A. and Wing, R.A.
TITLE A BAC End Sequencing Framework to Sequence the Rice Genome
JOURNAL Unpublished (1998)
COMMENT

Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Seq primer: TAATAGACTCACTATAGGG
 Class: BAC ends
 High quality sequence start: 17
 High quality sequence stop: 192.

FEATURES

source
 1. .614
 . /organism="Oryza sativa"
 /strain="Japonica"
 /cultivar="Nipponbare"

/note="Vector: pBelosBAC11; Site_1: HindIII; Site_2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

/db_xref="taxon:4530"
 /clone_lib="CUGI Rice BAC Library"
 /tissue_type="Leaf"
 /lab_host="E. coli DH10B"
 /lab_type="E. coli DH10B"

BASE COUNT 169 a 112 c 126 g 207 t
ORIGIN

Query Match 73.9%; Score 17; DB 28; Length 614;
 Best Local Similarity 100.0%; Pred. No. 2.26e-03;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 230 GCAAGGACAGGTTTCA 246
 |||||
Cp 23 gcaaggacagggtttca 7

RESULT 5
LOCUS MMTSG160X 194 bp RNA EST 22-JUN-1992
DEFINITION Mouse expressed sequence tag (EST) mRNA Tsg160x, mRNA sequence.
ACCESSION X61815
NID 954972
KEYWORDS EST; expressed sequence tag.
SOURCE house mouse.
ORGANISM Mus musculus

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae; Murinae; Mus.
 1 (bases 1 to 194)

REFERENCE
AUTHORS Hoog, C.
TITLE Direct Submission
JOURNAL Submitted (30-AUG-1991) C. Hoog, Dept of Mol Genetics, Karolinska Institutet, Box 60400, 10401 Stockholm, SWEDEN

REFERENCE

2 (bases 1 to 194)
AUTHORS Hoog, C.
TITLE Isolation of a large number of novel mammalian genes by a differential cDNA library screening strategy
JOURNAL Nucleic Acids Res. 19 (22), 6123-6127 (1991)
COMMENT 92066457

FEATURES Location/Qualifiers
 source
 1. .194
 /organism="Mus musculus"
 /strain="BALB/c"
 /db_xref="taxon:10090"
 /clone_lib="Lambda uni-ZAP II"

BASE COUNT 51 a 55 c 53 g 34 t
ORIGIN

Query Match 69.6%; Score 16; DB 20; Length 194;
 Best Local Similarity 90.0%; Pred. No. 3.94e-02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 26 GCAATGACAGGTTTCAGCC 45
 |||||
Cp 23 gcaaggacagggttcagcc 4

RESULT

6 AA732313 196 bp mRNA EST 06-JAN-1998
LOCUS nz91g02.s1 NCI-CGAP-GCBI Homo sapiens cDNA clone IMAGE:1302866
DEFINITION similar to contains Alu repetitive element; contains element PTR5 repetitive element ; mRNA sequence.

ACCESSION AA732313
NID 92752920
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
 Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 196)

REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,

Ph.D., Gerald Marti, M.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/db/rp/image/image.html

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 136.

FEATURES

source
 1. .196
 Location/Qualifiers

/organism="Homo sapiens"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTTTTT-3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 41 a 60 c 37 g 58 t
 ORIGIN

/db_xref="taxon:9606"
 /clone="IMAGE:1302866"
 /clone_lib="NCI_CGAP_GCB1"
 /tissue_type="germinal center B cell"
 /lab_host="DH10B"

Query Match 69.6%; Score 16; DB 11; Length 196;
 Best Local Similarity 86.4%; Pred. No. 3.94e-02;
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 20 GCAGGACAGGGTTTCACCATTA 41
 ||| ||||| ||||| |||
 Cp 23 gcaaggacaggggttcagccta 2

RESULT 7 AA990507 248 bp mRNA EST 02-JUN-1998
 LOCUS ua62gll.s1 Soares 2NbMT Mus musculus cDNA clone 1362116 3', mRNA
 DEFINITION sequence.
 ACCESSION AA990507
 NID G3175871
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 248)
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 TITLE The WashU-HMT Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT

Contact: Marra M/Mouse EST Project
 WashU-HMT Mouse EST Project
 Washington University School
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:895336
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 207.
 Location/Qualifiers

FEATURES
 source

1..248
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /note="Vector: p7T3D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5',
 TGTTACCAATCTGAAGTGGAGCGCGCGTGTGTGTGTGTGTGTGTGT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. RNA
 provided by Dr. Bertrand Jordan. Library went through two
 rounds of normalization, and was constructed by Bento
 Soares and M.Patima Bonaldo."
 /db_xref="taxon:10090"
 /clone="1362116"
 /clone_lib="Soares 2NbMT"
 /sex="male"
 /tissue_type="Thymus"
 /dev_stage="4 weeks"
 /lab_host="DH10B"

68 a 67 c 68 g 45 t

BASE COUNT
 ORIGIN

Query Match 69.6%; Score 16; DB 14; Length 248;
 Best Local Similarity 90.0%; Pred. No. 3.94e-02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 46 GCATGACAGGTTTCAGCC 65
 ||||| ||||| ||||| |||||
 Cp 23 gcaaggacaggggttcagcc 4

RESULT 8 A0106018 251 bp DNA GSS 27-AUG-1998
 LOCUS HS_3050_B1_H03_MR CIT Approved Human Genomic Sperm Library D Homo
 DEFINITION sapiens genomic clone Plate-3050 Col=5 Row=P, genomic survey
 sequence.
 ACCESSION A0106018
 NID G3481374
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 251)
 AUTHORS Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,
 Iraicoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.F.
 TITLE Construction of a Characterized Clone Resource for Genomic
 Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
 Tagged Connectors
 JOURNAL Unpublished (1997)
 COMMENT

Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Sequence Tagged Connector
 Plate: 3050 row: P column: 5
 Class: BAC ends
 High quality sequence stop: 251.
 Location/Qualifiers

FEATURES
 source

1..251
 /organism="Homo sapiens"
 /note="Organ: sperm; Vector: pBelobAC11; BAC Clones In
 E-Coli DH10B"
 /db_xref="taxon:9606"
 /clone="Plate-3050 Col=5 Row=P"
 /clone_lib="CIT Approved Human Genomic Sperm Library D"
 /sex="male"

BASE COUNT 45 a 66 c 46 g 94 t
 ORIGIN

Query Match 69.6%; Score 16; DB 28; Length 251;
 Best Local Similarity 86.4%; Pred. No. 3.94e-02;
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 182 GCAGGGACAGGGTTTCACCATTA 203
 ||| ||||| ||||| |||
 Cp 23 gcaaggacaggggttcagccta 2

RESULT 9 R57773 270 bp mRNA EST 02-MAY-1996
 LOCUS F5516 Fetal heart Homo sapiens cDNA clone F5516 5' end similar to
 DEFINITION DNA-binding protein TAXREB107, mRNA sequence.
 ACCESSION R57773
 NID 9827831
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
 Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 270)


```

/organism="Homo sapiens"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
(5'-TGTACCAATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTTTTTT-
3'). Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone_lib="IMAGE:1353885"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"

BASE COUNT      72 a      87 c      54 g      91 t
ORIGIN

Query Match      69.6%; Score 16; DB 14; Length 304;
Best Local Similarity 86.4%; Pred. No. 3.94e-02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 20 GCAGGGACAGGGTTTCCACCATA 41
|||||
Cp 23 gcaaggacaggttcagccta 2

RESULT 12
LOCUS      AA684131      320 bp      mRNA      EST      09-DEC-1997
DEFINITION      vm53b05.r1 Stratagene mouse Tcell 937311 Mus musculus cDNA clone
1001937 5', mRNA sequence.
ACCESSION      AA684131
NID      92670717
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryotae; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 320)
AUTHORS      Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilton,R. and
Waterston,R.
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:566153
TITLE      The WashU-HHMI Mouse EST Project
JOURNAL      Unpublished (1996)
COMMENT      Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:566153
Seq primer: -28ml3 revl ET from Amersham
High quality sequence stop: 286.
Location/Qualifiers
1. .320
/organism="Mus musculus"
/notes="Organ: blood; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. M30 CD4+ cells. Average insert size: 1.0 kb;
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGCGCAGG
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3"
/db_xref="taxon:10090"
/clone="1001937"

/organism="Homo sapiens"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
(5'-TGTACCAATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTTTTTT-
3'). Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone_lib="IMAGE:1353885"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"

BASE COUNT      72 a      87 c      54 g      91 t
ORIGIN

Query Match      69.6%; Score 16; DB 14; Length 304;
Best Local Similarity 86.4%; Pred. No. 3.94e-02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 20 GCAGGGACAGGGTTTCCACCATA 41
|||||
Cp 23 gcaaggacaggttcagccta 2

RESULT 13
LOCUS      W18476      329 bp      mRNA      EST      10-SEP-1996
DEFINITION      mb54h09.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 333281 5'
similar to gb:250159 M.musculus mRNA for Sull (MOUSE);, mRNA
sequence.
ACCESSION      W18476
NID      91294200
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryotae; Chordata; Vertebrata; Mammalia; Eutheria; Muridae; Murinae;
Mus.
REFERENCE      1 (bases 1 to 329)
AUTHORS      Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilton,R. and
Waterston,R.
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:214681
TITLE      The WashU-HHMI Mouse EST Project
JOURNAL      Unpublished (1996)
COMMENT      Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:214681
Seq primer: ETPrimer
High quality sequence stop: 293.
Location/Qualifiers
1. .329
/organism="Mus musculus"
/notes="Vector: pT73D (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer 15'
TGTACCAATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTTTTTT 3',
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector:
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. RNA was kindly provided by
Dr. Minoru KO (Wayne State University)."
/db_xref="taxon:10090"
/clone="333281"
/clone_lib="Soares mouse p3NMF19.5"
/dev_stages="19.5 dpc total fetus"
/lab_host="DH10B (ampicillin resistant)"
<1. .329

BASE COUNT      93 a      84 c      77 g      75 t
ORIGIN

```

```

/clone_lib="Stratagene mouse Tcell 937311"
/tissue_type="Tcell"
/dev_stage="M30 CD4+ cells"
/lab_host="SOLR (kanamycin resistant)"

BASE COUNT      89 a      81 c      81 g      69 t
ORIGIN

Query Match      69.6%; Score 16; DB 11; Length 320;
Best Local Similarity 90.0%; Pred. No. 3.94e-02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 120 GCAATGACAGGTTTCAGCC 139
|||||
Cp 23 gcaaggacaggttcagcc 4

RESULT 13
LOCUS      W18476      329 bp      mRNA      EST      10-SEP-1996
DEFINITION      mb54h09.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 333281 5'
similar to gb:250159 M.musculus mRNA for Sull (MOUSE);, mRNA
sequence.
ACCESSION      W18476
NID      91294200
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryotae; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.
REFERENCE      1 (bases 1 to 329)
AUTHORS      Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilton,R. and
Waterston,R.
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:214681
Seq primer: ETPrimer
High quality sequence stop: 293.
Location/Qualifiers
1. .329
/organism="Mus musculus"
/notes="Vector: pT73D (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer 15'
TGTACCAATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTTTTTT 3',
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector:
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. RNA was kindly provided by
Dr. Minoru KO (Wayne State University)."
/db_xref="taxon:10090"
/clone="333281"
/clone_lib="Soares mouse p3NMF19.5"
/dev_stages="19.5 dpc total fetus"
/lab_host="DH10B (ampicillin resistant)"
<1. .329

BASE COUNT      93 a      84 c      77 g      75 t
ORIGIN

```



```

Query Match          69.6%; Score 16; DB 21; Length 329;
Best Local Similarity 90.0%; Pred. No. 3.94e-02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 170 GCAATGACAAAGGTTTCAGCC 189
      |||| |||| |||| |||| ||||
Cp 23 gcaagacaggggttcagcc 4

RESULT 14
LOCUS      AA065583      342 bp      mRNA      EST      03-FEB-1997
DEFINITION m173g05.r1 Stratagene mouse kidney (#937315) Mus musculus CDNA
ACCESSION  AA065583
NID        91563047
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.
REFERENCE  1 (bases 1 to 342)
AUTHORS    Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
TITLE      The WashU-HHMI Mouse EST Project
JOURNAL    Unpublished (1996)
COMMENT    Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:311536
            Seq primer: -28ml3 rev1 ET from Amersham.

FEATURES   Location/Qualifiers
            source
            1..342
            /organism="Mus musculus"
            /strain="C57/BL6"
            /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
            XhoI; cloned unidirectionally. Primer: Oligo dt. Average
            insert size: 1.0 kb; Uni-ZAP XR Vector: -5' adaptor
            sequence: 5' GAATTCGCGCAGAG, 3' -3' adaptor sequence: 5'
            CTTGAGTTTCTTTTCTTTTCTTTT 3'."
            /db_xref="taxon:10090"
            /clone="517688"
            /clone_lib="Stratagene mouse kidney (#937315)"
            /sex="females"
            /dev_stage="4 weeks"
            /lab_host="SOLR (kanamycin resistant)"
            <1..>342
BASE COUNT  89 a      86 c      96 g      71 t
ORIGIN

Query Match          69.6%; Score 16; DB 23; Length 342;
Best Local Similarity 90.0%; Pred. No. 3.94e-02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 40 GCAATGACAAAGGTTTCAGCC 59
      |||| |||| |||| |||| ||||
Cp 23 gcaagacaggggttcagcc 4

RESULT 15
LOCUS      W00972      361 bp      mRNA      EST      18-APR-1996
DEFINITION za55d07.r1 Soares fetal liver spleen lNFLS Homo sapiens CDNA clone
            296461 5' similar to contains element L1 repetitive element; , mRNA
            sequence.

Accession
NID        91272951
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 361)
AUTHORS    Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and
            Wilson,R.
TITLE      The WashU-Merck EST Project
JOURNAL    Unpublished (1995)
COMMENT    Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: mob.REGA+ET
            High quality sequence stop: 217.

FEATURES   Location/Qualifiers
            source
            1..361
            /organism="Homo sapiens"
            /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
            with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
            1st strand cDNA was primed with a Pac I - oligo(dt) primer
            [5' AACTGGGAAGTAATAATTAAGATCTTTTTTTTTTTTTTTT 3'],
            double-stranded cDNA was ligated to Eco RI adaptors
            (Pharmacia), digested with Pac I and cloned into the Pac I
            and Eco RI sites of the modified pT7T3 vector. Library
            went through one round of normalization. Library
            constructed by Bento Soares and M.Fatima Bonaldo."
            /db_xref="taxon:9606"
            /clone="296461"
            /clone_lib="Soares fetal liver spleen lNFLS"
            /sex="male"
            /dev_stage="20 week-post conception fetus"
            /lab_host="DH10B (ampicillin resistant)"
            <1..>361
BASE COUNT  103 a      85 c      81 g      90 t      1 others
ORIGIN

Query Match          69.6%; Score 16; DB 20; Length 361;
Best Local Similarity 94.4%; Pred. No. 3.94e-02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 329 AGGCTGAACCTGTCCT 346
      ||| |||| |||| |||| ||||
Cp 3 aggtgaaacctgtcct 20

Search completed: Thu Apr 1 05:30:28 1999
Job time : 313 secs.

```

THIS PAGE BLANK (USPTO)